Assignment 3 – Quantitative Analysis of Australian's Interest and Engagement with Science

Usability Engineering 2019 Semester 1

Todd Johnston – s3660590

Contents

Introduction

The purpose of this study was to Quantitatively analyse the provided data from Australian participants during National Science Week, from 13 to 21 August 2011[1]. The data was collected in the form of a survey which contained multiple types of data including binary, non-binary and descriptive datasets. The data received was incomplete and as a result some data cleaning was performed to attempt to increase the integrity of the results. As the data is relatively old and erroneous, all results should be taken into perspective and may not be fully indicative of the population of Australia. To analyse the data 2 research questions were provided and another 2 created to comprise the basis of this report. These questions were answered in a both quantitative and qualitative manner and the results formatted by each individual question.

Research Questions

Q1. We are interested in the demographics of the people in Australia who are most interested in Science. Can you find where they live, which age group do they belong and what is their gender?

For this first question the data was separated so that only entries containing the value "Very interested" for the interest in science attribute were represented. This was to ensure that only those users with the most interest in science were being analysed. The data was then cleaned so as to emit those users who had not listed a location, as these users were seen as irrelevant to the question.

Q2: We would like you to consider those who have an education in Science and find out whether or not they are employed in a Science-based job.

To capture those users who "Have an education in Science" the data was once again selectively measured so that only users with a tertiary education in Science were considered. This was chosen to be the limiting attribute as high number of users likely took compulsory or elective high school science subjects that would not necessarily convey to employability. Tertiary education on the other

hand is considered much more of a necessity and as a result would likely reflect the job market at the time of the study more accurately.

Q3. Is there a relation between the level of science interest and talking about science with friends and family?

The purpose of this research question was to find the level of which user's science interest affected their communication on the subject with their friends and family. The sample size(n = 3074) for this question used whole survey for the population as the data was found to be complete for all the criteria which was essential to the question.

Q4. For what uses have users found Science most helpful?

The intention behind this question was to uncover the ways in which users have found uses for their general science knowledge or skills and how often these skills are used. Once again, the entire sample size was used as the binary data was found to be complete for all attributes being analysed. The original survey posed four questions about the use of the data science in daily lives: Whether users used science at work, for their hobbies or interests, to help solve a particular problem, for help with purchasing something and to develop their understanding of a health issue.

Analysis

Q1.

As seen in figure 1, participants came in varying amounts from the different states and territories, with those along the east coast providing the most users. From this we can begin to draw an idea about the location of our users.
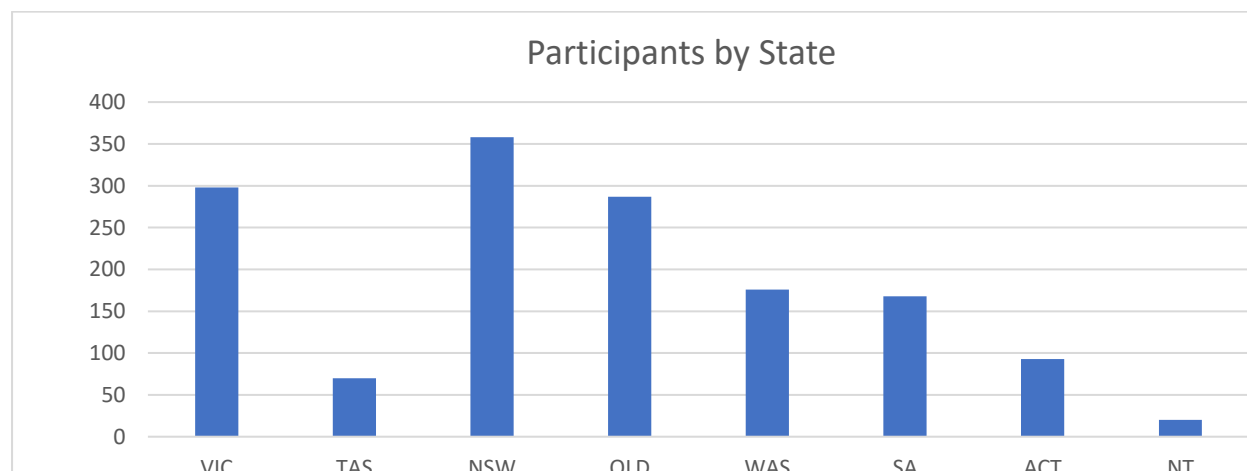

Figure 1 - Number of participants per state

In terms of age, both the median and the mean (accounting for rounding) of our sample population was 42 years of age. There was some variance across the regions when comparing age groups (figure 2.) but the states generally followed the same trend with the age groups of 18-24 years old's and 65+ year old's being significantly lower than all others. This indicates that there was either an under-representation in these age groups or these groups simply have less interest in science matters. This suggests that the survey was accidentally targeted towards those age groups which are over-represented.
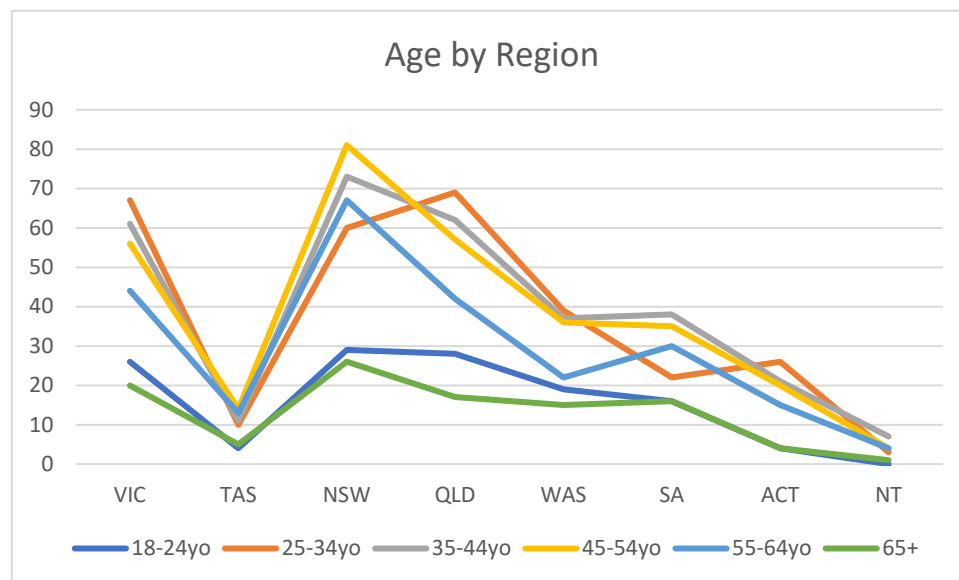

Figure 2 - Age group bands compares across the regions

In terms of gender, across all states females were represented in higher numbers than the male demographic. This was particularly evident in states such as New South Wales and Victoria where the number of users was higher and the rate at which females appeared was nearly double that of males. Overall the data (Figure 3.) shows that of those users who are most interested in science, there is a 60.7% chance that they are female, signifying that females have a higher interest in science across the general sample population.

	VIC	TAS	NSW	QLD	WA	SA	ACT	NT	Total	
Males	113	30	124	117	75	64	46	8	577	39.3%
Females	185	40	234	170	101	104	47	12	893	60.7%

Figure 3 – User's Gender By State

From this data we can draw a sample of mode user represented by this sample size, a female 45-54-year-old from the region of NSW. This modal user can be used to reference against potential bias's in other questions, such as in the case of whether or not a user has a job in science, as it is more likely for our modal user to have employment as opposed to a 18-24 year old or 65+ year old.

Q2.

The total number of users in our new sample, those with a tertiary education, is 881. Of these users 363, or 41.2% (Figure 4.), answered that they currently held a job in a science-based field. Interestingly, a further 21% of users indicated that whilst they no longer held a job in the science industry, they were previously employed in the field. As a result this indicates that those users who pursue further education in science have just over a 60% chance to gain employment in the field. This statistic indicates great opportunity for those who are interested in science to be able to pursue a career in line with their interests. Interestingly, of the original sample size of users who are very interested in science, the number of users with jobs in the industry sits square at 25%, begging the question of whether or not there is the desire to convert the interest into further studies and then a career.
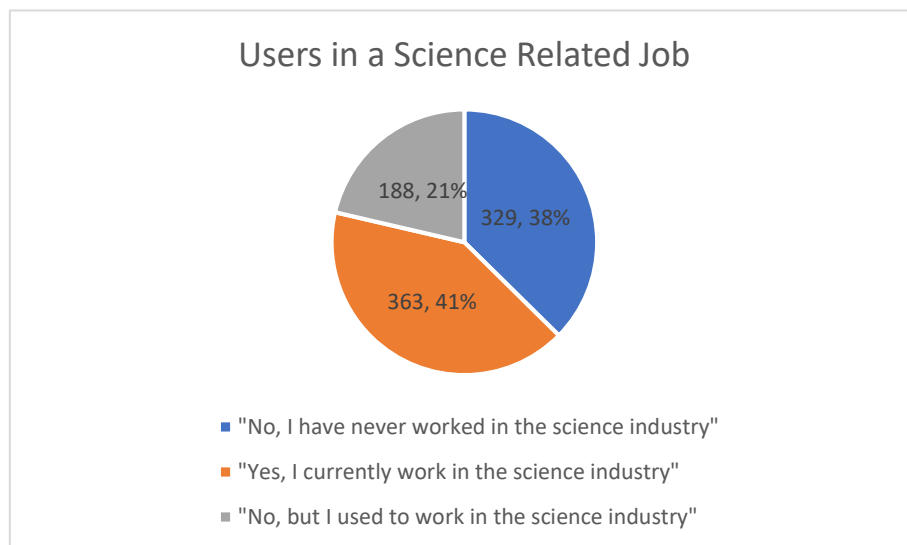

Users in a Science Related Job

- 329, 38%
- 363, 41%
- 188, 21%

- ■ "No, I have never worked in the science industry"
- ■ "Yes, I currently work in the science industry"
- ■ "No, but I used to work in the science industry"

Figure 4 - Users in Science Related Jobs

Q3.

The results showed an extremely heavy correlation at the top end, with those being "Very Interested" or "Quite Interested" in science being much more likely to discuss it with friends and/or family, as shown by figure 5. This correlation could be somewhat expected but the extent to which it is followed is somewhat surprising, even those who are "Neither interested nor uninterested" in science barely discuss the topic, suggesting that it is not a topic which creates great interest outside of its community. For those who are interested however science is clearly discussed at high rates

with both friends and family, implying that users with curiosity about the topic tend to share it with those close to them.

Figure 5 - How Does Interest in Science Relate to Frequency of Science Based Conversation?

Relation Between Interest In Science and Sharing of Science Info
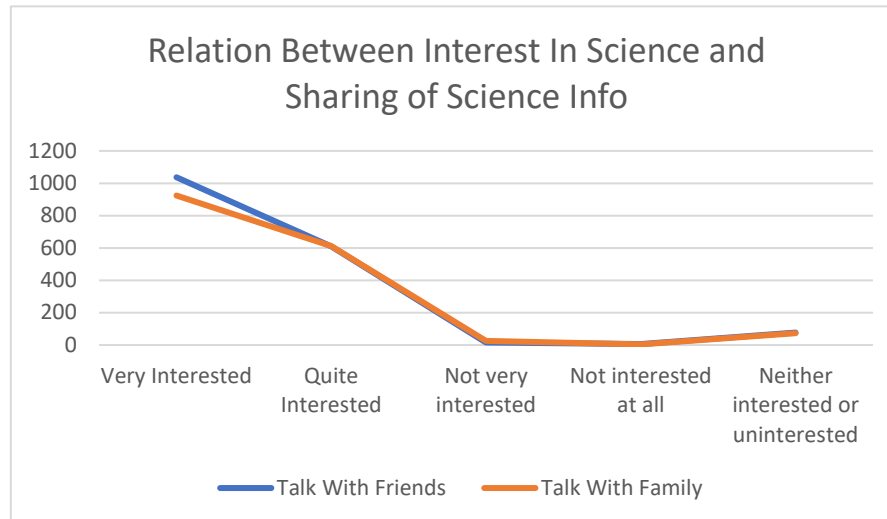
Q4.

The results of this analysis revealed that many users have found a use for science in some area or another, showing the importance of science in numerous participants day to day lives. As seen in Figure 6., science is most commonly used to help develop a user's understanding of a health issue. This figure could be influenced by the correlation between the science and medical fields perhaps biasing the resultant data. The same can be said for the use of science to help purchase items, as science knowledge could only applied to field related products thus skewing the results. Also important to note is the apparent vagueness of the "Use it to help me solve a particular problem" attribute. As this attribute does not clearly describe its criteria it has a very broad range thus inflating its number of positive responses. The decision was made however to leave it in the analysis as its result was not a clear outlier and the original survey question may have been more descriptive on its measures.
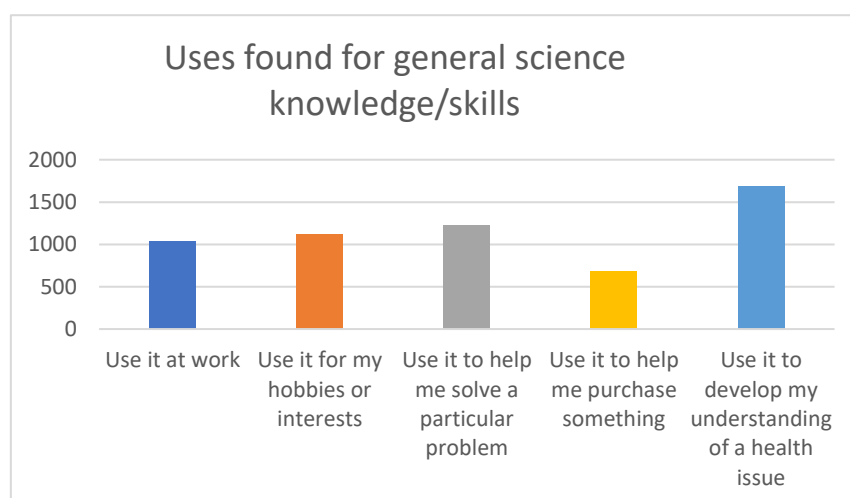

Uses found for general science knowledge/skills

Figure 6 - Uses for Science in User's Life

Limitations

The data that was used to make this analysis was not perfect, with several entries missing data or having unfiltered answers. Some questions from the survey were even unlabelled, as a result the integrity of the original data has to be called into question as the number of errors in the collection was rather high. The sample size for some questions, such as in Q1 and Q2, had to be reduced by a small amount to omit those entries which contained no data so as not to skew the calculations. As a result, a smaller sample population was taken which may have resulted in some bias in the results.

Cleaning the data

For some data entries the state in which they were located was not recorded. Many of these entries had stated that they were very interested in science, leaving a dilemma as to get rid of the entries would result in a loss of some data. In the end I decided against including them with the analysed data as the location was to relevant to the questions propositioned by the report

Conclusion

The outcome of the analysis performed on this data allowed us many insights into they way in which science plays a role in the life of Australians. From the survey we were able to conclude multiple factors about individual users and their demographic details such as gender, age and location within Australia. The data also provided us with an understanding as to how tertiary education can play a role in pursuing a career in the field of science and its relevant industries. The third research question also shed light on communication and fact sharing amongst the participants and its correlation between the users varying interest levels on the topic. The final segment of the analysis also revealed more about the daily uses of science and the relation between differing fields influencing participants need for science. Conclusively, the analysis performed on this data was done so with consideration to both the population size and data integrity used in quantitative and qualitative investigation methods and as such this report should be referenced with the same consideration.

References

1. (20/05/19 – 26/05/19) Raw Data sourced from Victorian government Department of State Development, Business and Innovation (DSDBI). Available at <https://www.data.vic.gov.au/data/dataset/australians-interest-and-engagement-with-science/resource/dc730e2d-f6af-47c9-844c-9cd6299ffb8e>