final report

Project code: L.EQT.1804

Prepared by: Garth Tarr
                            Polkinghorne's Pty Ltd

Date published: 3 May 2018

Statistical Evaluation and Workshop (Analytics for Industry)

Milestone 2: Delivery of one 5 day statistical workshop

Abstract

The meat processing and pastoral industries are rapidly entering the digital age and are accumulating large and complex databases from properties, feedlot and processing systems. The challenge is to integrate this data into the business using appropriate analyses to create value for the business and flow on benefits to industry.

Given that much of the data is sensitive it is most effective that analysis be done in-house, so the data handling skills of the relevant staff need to be advanced in order to handle the large volume of data which will be delivered back to the processing and production sectors.

Currently there are a number of courses available commercially however these are not targeted to agriculture and tend to be very expensive which deters companies from investing in these courses. Hence MSA/MLA saw a need to develop and run a course targeted to the data input and problems specific to the processing and pastoral sectors. It would provide a mechanism for current staff to upskill and better service company needs, with a longer-term outcome being that such courses will be able to be run in the future which will develop a pool of users with relevant skills who are focused on agricultural problems.

This project developed and delivered statistical training to company staff on a cost recovery basis to develop the data handling/analysis skills of the meat processing and pastoral industries. The course is aimed at introducing students to the possibilities that are available in data analysis and visualisation of data by providing them with basic R skills.

Table of contents

1   Background

The meat processing and pastoral industries are rapidly entering the digital age and are accumulating large and complex data bases from properties, feedlot and processing systems.

These industries are not well prepared for this onslaught of data in terms of staff with the appropriate skills and tools to collate, prepare and analyse these complex databases.

To date the pastoral and meat processing companies have employed graduates with a rudimentary knowledge of data handling (largely excel, with some basic knowledge of other statistical packages) and a passion for agriculture. Data handling skills for these staff need to be advanced if they are to handle the large volume of satellite, pasture, animal performance and carcass data which will be delivered back to the processing and production sectors. The challenge is to integrate this data into the business using appropriate analyses to create value for the business.

In addition, the recent push for objective carcass measurement (OCM) will accelerate the need for data processing skills to be developed in processing companies to underpin development of value based training systems. Both the OCM and lamb supply chain projects will have a need for supply chain officers to have advanced data handling skills to help processing works undertake analyses in-house.

Given that much of the data is sensitive it is critical that these analyses be done in-house where the
- objectives in the analyses be more clearly articulated
- there is continuity in the types of analyses being performed

It is much more cost efficient to undertake these analyses in-house than sub-contracting to external providers.

2   Project objectives

To deliver statistical training to company staff to develop the data handling/analysis skills of the meat processing and pastoral industries. The course is aimed at introducing students to the possibilities that are available in data analysis and visualisation of data by providing them with basic R skills.

Whilst the data problems being presented to staff are diverse, the course will achieve several things; firstly, it will show them how to address a statistical problem and break it up into logical steps. The R packages introduced to the students will then assist them to convert these steps into statistical code, running the code and then producing an output.

The longer-term advantage of running such courses is that it develops a pool of users that have focused on agricultural problems. This like-minded group will provide the bases for a user group that can meet regularly to share problems and solutions. It will upskill the group so that in addition to sharing problems they will begin to use the web resources to further refine their skills. This will not occur without delivering basic courses to introduce students to the packages and get them started. The course would be run on a cost recovery basis and specifically focus on handling mapping, pasture, live animal, carcass quality and carcass yield data. There are a number of courses that are provided by the commercial sectors however these are not targeted to the agricultural applications and tend to be very expensive which deters companies from investing in these courses. The logic for MSA/MLA to run the course would be that it was targeted to the data input and problems which are

specific to the processing and pastoral sectors. It would provide a mechanism for current staff to upskill and better service company needs.

The course frequency would be determined by the demand from industry. Obviously, there would be a call for more frequent courses initially and then perhaps one or at most two a year.

There are a number of software packages are capable of handling complex data sets. These include "R", "SAS", "SPSS", "GENSTAT" and "ASREML" etc. Whilst all these programs tend to have a similar capability in terms of the statistical analyses they undertake, they do differ greatly in both the visual quality of the outputs and the initial cost of the programs.

Among R's many features is the capability to output high quality static plots and figures. Increasingly, there are many interactive visualisations that can also be created with R, including geospatial mapping applications and web-based interfaces so that end-users can query data sets and analysis without needing to know any R programming.

3   Methodology

One 5 day statistical workshop

The 5 day workshop was run from 4-6 April and 19-20 April at the MLA Brisbane office. There were 7 participants taught by Dr Garth Tarr (University of Sydney) and Mr Kevin Wang (University of Sydney). A short biography for both presenters can be found in the appendix.
A course webpage (https://garthtarr.github.io/meatR) has been developed to host the course materials and provides a reference portal for past participants. It will continue to be added to and refined over time. The majority of course preparation work is in building, updating and adding to the course webpage.

4   Results

At the end of the 5 day workshop, the participants each gave presentations an analyses they'd performed in R on their own data sets. Example analyses include:

- Anomaly detection in weather data with a view to predicting dark cutting
- Looking at connecting databases to interact directly with R
- Understanding measurement variability and reliability over time with calibration interventions
- Modelling eating quality with linear mixed models
- Amalgamating data sets from various sources

Image 1: Workshop participants

Feedback

It's essential to collect participant feedback to help improve the structure and delivery of the workshop over time. A survey was created to collect participant feedback and participants were strongly encouraged to provide their feedback on the final day. The feedback form can be found here

https://docs.google.com/forms/d/e/1FAIpQLScVkJ9yWGbLLG4_nCxR1f4IZAX156wcdJbNHL8oUsdfk4YLLQ/viewform

All 7 participants responded to the survey. We had 100% agreement (or strong agreement) to the following statements:

About the presenters:
- The presenters communicated the information clearly
- The presenters made the subject matter compelling
- The presenters were able to answer questions
- The presenters were friendly and approachable

Unedited comments:

Practical and fluidly adaptable
Garth and Kevin were both excellent teachers. Were both very helpful, informative and friendly.
Both presenters had a great way of explaining the concepts of R, and worked us through the example data in a way that made it easy to understand the concepts. Furthermore, their friendly nature and sense of humour made the learning experience even more enjoyable.

The presenters were very informative and willing to assist and solve problems. Having few people in the class, meant the presenters were able to spend more time with individuals which was good.

This course was great. Not only was the information presented and explained in an easy to understand way but the presenters were always approachable to ask questions.

We also had 100% agreement (or strong agreement) to the following statements

About the material:
- The workshop was relevant to me
- The workshop was interesting
- The workshop met my purpose in attending
- I want to tell others about what was presented
- The website resources were useful
- The opportunity to work on my own data was useful

Comments:

was able to apply the theory to my problem

Very well structured course and delivery of content.

The material complemented the concepts that we were learning. Furthermore, we used data that were relevant to our industry.

The material used in the workshop was very industry relevant, which was extremely useful. I feel it made it more engaging and easy to understand what we wanted to do with the data, having the end outcome in mind.

This material and content presented in this course was relevant and presented in a way that the audience was able to relate and understand the topic. This isn't always the way with stats, so thank you!! It was great that you have the ability to work on our own data so that we can really understand and apply the course content.

About the workshop in general:
- The duration of the workshop was appropriate (1 Neutral, 6 agree or strongly agree)
- The workshop was well organised (100% agree or strongly agree)
- I would recommend this workshop to others (100% agree or strongly agree)

Comments:

I was very in-experienced in R but was well catered for amounst other who did

Fantastic - would like to take part in the follow up w/s

As always, the workshop could have been longer. There was some disruption earlier in the course around room bookings that made it a bit hard to learn.

I think the timing of the workshop was perfect. Three days, ten days to work on your own project, then two days at the end. I feel this was ample time to digest the information without it being to full on at once and enough time to make progress on your own data.

This was a great course. I was looking forward to learning about R and I have learnt a lot throughout the course. The additional tricks and explanation about material was really helpful.

What were the best aspects of the workshop?

Applicable

Presenter was very adaptable to the scope dependant on the continued feedback as we progressed

Practical application of your own data with assistance on hand. Learn other aspects of R which may be used in the future

Being someone that has never used R the structure and flowthrough in the delivery of content was good and easily understood.

Learning from the ground up and then having the ability to apply these learnings to data that I am used to working with.

Having like minded people with similar aims in mind.

Starting from the basics and building up. The way the information was delivered.

What aspects of the workshop need improvement?

Pace of learning with the pace of teaching sometimes got seperated

Nothing really much to improve on.
Maybe slow down a little when working through examples, or tell everyone not to type and copy, then give an opportunity to copy once it is explained following.

Potentially the last day, we didn't get as much done? But was good to see presentations, the work everyone had done.

Some parts could be run through slower so that you can catch up.

Please give a couple of concrete examples of how you will be able to apply what was learnt in the workshop to improve your organisation's decision making abilities and drive change.

Modelling pasture for pasture values

Build a web interfaced dashboard using Shiny to share info with employees

No longer need to rely on others for models/graphs for PhD, work or any other use

Majority of my work requires me to be able to process and eventually analyse raw data. I have only previously used excel, spss and sas and being confident in one program for all of my input, data manipulation and outputs will save me infinitely on time.

I will be using the learnings from this workshop to develop a simple way for our team to bring in data directly from the database and manipulate it to get meaningful answers to give to our stakeholders.
I will also use this workshop to work on random data sets and projects that will streamline processes.

I learnt so much in this course, I can't even begin to explain. I have the confidence to go out and learn more about R and better ways to analyse data. I believe I will be of more use to the company now and be able to drive change.

I will be able to use R studio. I can now use the program to clean, tidy, and run models.

Do you have any suggestions about the best way to continue to support and grow your analytics skills in R?

Opportunity to work in depth on own data with council

Continued practice. Hands on courses every yr (or even a day or two each year)

More workshops/annual upskill

I think one more follow up, maybe just 2-3 days would be good. Just to see how far everyone's come, and learn a few more advanced skills on top, just to finish off.

Online groups and conferences.

Any other comments?

Thank you very much

no

Great course, got a lot out of it.

THANK YOU!

5 Conclusions/recommendations

Recommendation 1
Continue with the 5 day format for initial Analytics for Industry training workshops. The dates should be finalised months in advance and advertised widely. The recommended number of participants remains 7-10. If the course is held in Brisbane again, the boardroom at the MLA offices was a perfect venue.

Recommendation 2
There seems to be demand for annual boot camps or refresher courses to build on and consolidate the techniques learnt in the initial workshop. Continue with the 2 day follow up.

6 Key messages

- The 5 day workshop was a clear success.
- Fix course dates months before hand and advertise widely.
- Move towards a "data retreat" model for annual refresher courses/boot camps.

7 Bibliography

Tarr (2017). Analytics for Industry. Web resource. https://garthtarr.github.io/meatR/

8 Appendix

8.1 Short biographies of the presenters

Dr Garth Tarr (University of Newcastle)
Garth is a lecturer in Statistics and Data Science at the University of Sydney. He received his PhD in Mathematical Statistics from the University of Sydney and has held positions at the Australian National University and University of Newcastle. Garth is an expert R user, he has developed three R packages and contributed to a number of others. His diverse interests include data visualisation, meat science, robust statistics, model selection, econometric modelling (including value based marketing), educational research and biostatistics.

Kevin Wang (University of Sydney)
Kevin is currently a PhD candidate and Postgraduate Teaching Fellow in the School of Mathematics and Statistics at the University of Sydney. Kevin's main research area is in statistical bioinformatics and is developing novel methods brought forward by high dimensional biomedical data. A central focus of his current research focuses on the increasingly popular boutique array platform and its application both as a validation platform for biomarkers for patients in melanoma studies. His other interests include data visualisation, statistical computing, model selection and biostatistics.

8.2 Full survey results

8.2.1 About the presenters

Timestamp	ABOUT THE PRESENTERS [The presenters communicated the information clearly]	ABOUT THE PRESENTERS [The presenters made the subject matter compelling]	ABOUT THE PRESENTERS [The presenters were able to answer questions]	ABOUT THE PRESENTERS [The presenters were friendly and approachable]	ABOUT THE PRESENTERS
20/04/2018 15:36:47	Agree	Agree	Agree	Strongly Agree	Practical and fluidly adaptable
20/04/2018 15:38:48	Agree	Agree	Strongly Agree	Strongly Agree	
20/04/2018 15:41:22	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	
20/04/2018 15:41:48	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Garth and Kevin were both excellent teachers. Were both very helpful, informative and friendly.
20/04/2018 15:42:13	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Both presenters had a great way of explaining the concepts of R, and worked us through the example data in a way that made it easy to understand the concepts. Furthermore, their friendly nature and sense of humour made the learning experience even more enjoyable.
20/04/2018 15:42:30	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	The presenters were very informative and willing to assist and solve problems. Having few people in the class, meant the presenters were able to spend more time with individuals which was good.
20/04/2018 15:42:57	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	This course was great. Not only was the information presented and explained in an easy to understand way but the presenters were always approachable to ask questions.

8.2.2 About the material

ABOUT THE MATERIAL [The workshop was relevant to me]	ABOUT THE MATERIAL [The workshop was interesting]	ABOUT THE MATERIAL [The workshop met my purpose in attending]	ABOUT THE MATERIAL [I want to tell others about what was presented]	ABOUT THE MATERIAL [The website resources were useful]	ABOUT THE MATERIAL [The opportunity to work on my own data was useful]	ABOUT THE MATERIAL
Agree	Agree	Agree	Agree	Agree	Agree	was able to apply the theory to my problem
Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Strongly Agree	
Strongly Agree	Strongly Agree	Agree	Strongly Agree	Agree	Strongly Agree	
Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Very well structured course and delivery of content.
Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Strongly Agree	The material complemented the concepts that we were learning. Furthermore, we used data that were relevant to our industry.
Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	The material used in the workshop was very industry relevant, which was extremely useful. I feel it made it more engaging and easy to understand what we wanted to do with the data, having the end outcome in mind.
Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	This material and content presented in this course was relevant and presented in a way that the audience was able to relate and understand the topic. This isn't always the way with stats, so thank you!! It was great that you have the ability to work on our own data so that we can really understand and apply the course content.

8.2.3 About the workshop in general

ABOUT THE WORKSHOP IN GENERAL [The duration of the workshop was appropriate]	ABOUT THE WORKSHOP IN GENERAL [The workshop was well organised]	ABOUT THE WORKSHOP IN GENERAL [I would recommend this workshop to others]	ABOUT THE WORKSHOP IN GENERAL
Neutral	Agree	Agree	
Agree	Strongly Agree	Strongly Agree	I was very in-experienced in R but was well catered for amounst other who did
Agree	Strongly Agree	Strongly Agree	Fantastic - would like to take part in the follow up w/s
Agree	Agree	Strongly Agree	
Agree	Agree	Strongly Agree	As always, the workshop could have been longer. There was some disruption earlier in the course around room bookings that made it a bit hard to learn.
Strongly Agree	Strongly Agree	Strongly Agree	I think the timing of the workshop was perfect. Three days, ten days to work on your own project, then two days at the end. I feel this was ample time to digest the information without it being to full on at once and enough time to make progress on your own data.
Agree	Strongly Agree	Strongly Agree	This was a great course. I was looking forward to learning about R and I have learnt a lot throughout the course. The additional tricks and explanation about material was really helpful.

8.2.4 Other comments

What were the best aspects of the workshop?	What aspects of the workshop need improvement?	Please give a couple of concrete examples of how you will be able to apply what was learnt in the workshop to improve your organisation's decision making abilities and drive change.	Do you have any suggestions about the best way to continue to support and grow your analytics skills in R?	Any other comments?
Applicable	Pace of learning with the pace of teaching sometimes got seperated	Modelling pasture for pasture values	Opportunity to work in depth on own data with council	
Presenter was very adaptable to the scope dependant on the continued feedback as we progressed		Build a web interfaced dashboard using Shiny to share info with employees		Thank you very much
Practical application of your own data with assistance on hand. Learn other aspects of R which may be used in the future	Nothing really much to improve on. Maybe slow down a little when working through examples, or tell everyone not to type and copy, then give an opportunity to copy once it is explained following.	No longer need to rely on others for models/graphs for PhD, work or any other use	Continued practice. Hands on courses every yr (or even a day or two each year)	
Being someone that has never used R the structure and flowthrough in the delivery of content was good and easily understood.		Majority of my work requires me to be able to process and eventually analyse raw data. I have only previously used excel, spss and sas and being confident in one program for all of my input, data manipulation and outputs will save me infinitely on time.		
Learning from the ground up and then having the ability to apply these learnings to data that I am used to working with.	NA	I will be using the learnings from this workshop to develop a simple way for our team to bring in data directly from the database and manipulate it to get meaningful answers to give to our stakeholders. I will also use this workshop to work on	More workshops/annual upskill	no

		random data sets and projects that will streamline processes.		
Having like minded people with similar aims in mind.	Potentially the last day, we didn't get as much done? But was good to see presentations, the work everyone had done.	I learnt so much in this course, I can't even begin to explain. I have the confidence to go out and learn more about R and better ways to analyse data. I believe I will be of more use to the company now and be able to drive change.	I think one more follow up, maybe just 2-3 days would be good. Just to see how far everyone's come, and learn a few more advanced skills on top, just to finish off.	Great course, got a lot out of it.
Starting from the basics and building up. The way the information was delivered.	Some parts could be run through slower so that you can catch up.	I will be able to use R studio. I can now use the program to clean, tidy, and run models.	Online groups and conferences.	THANK YOU!