

final report

Project code: L.EQT.1902
Prepared by: Garth Tarr
Polkinghorne's Pty Ltd
Date published: 29 September 2019

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

Analytics for Industry Course and Statistical Analysis Consultancy Agreement

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

This publication is published by Meat & Livestock Australia Limited ABN 39 081 678 364 (MLA). Care is taken to ensure the accuracy of the information contained in this publication. However MLA cannot accept responsibility for the accuracy or completeness of the information or opinions contained in the publication. You should make your own enquiries before making decisions concerning your interests. Reproduction in whole or in part of this publication is prohibited without prior written consent of MLA.

Abstract

This agreement has lead to the development of Shiny Apps for Heat Mapping, MSA Producer Awards, Beef Eating Quality Insights Report, BDO Toolbox as well as any other analytical capabilities for MSA funded research.

It has also delivered statistical training to company staff to develop the data handling/analysis skills of the meat processing and pastoral industries. The workshop 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.

Executive Summary

The primary outcome of this work is the development of the Business Development Office (BDO) toolbox, a web application built on the R statistical programming language that enables BDOs to process and analyse data efficiently and accurately, to extract insights and support the industry. It has been developed in close collaboration with MSA staff and is targeted to suit their needs.

We have also continued to support MSA through ad-hoc statistical support and the delivery of another Analytics for Industry workshop which was very well received by attendees. Contracts like these ensure that the MSA team has on demand access to cutting edge statistical support.

Table of Contents

1	Project Objectives	5
2	Results.....	5
2.1	MSA staff support.....	5
2.2	MSAapps R package	5
2.2.1	gcheck app	6
2.2.2	BDO toolbox	6
2.3	Workshop	6
2.4	Pathways attendance	7
3	Conclusions/Recommendations.....	7
4	Appendix.....	8
4.1	Workshop feedback.....	8
4.2	MSA Apps development timeline	11
4.2.1	MSAapps 0.2 [2019-06-25].....	11
4.2.2	MSAapps 0.1.9999 [2019-02-18]	11
4.2.3	MSAapps 0.1.9999 [2019-01-24]	11
4.2.4	MSAapps 0.1.9999 [2018-01-21]	11
4.2.5	MSAapps 0.1.9999 [2019-01-21]	11
4.2.6	MSAapps 0.1.9999 [2019-01-10]	11
4.2.7	MSAapps 0.1.9999 [2018-01-08]	11
4.2.8	MSAapps 0.1.999 [2018-01-08].....	12
4.2.9	MSAapps 0.1 [2018-01-06]	12

1 Project Objectives

The Consultant will achieve the following objective(s) to MLA's reasonable satisfaction:

- 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.
- Work with the MSA team to develop Shiny Apps for Heat Mapping, MSA Producer Awards, Beef Eating Quality Insights Report, BDO Toolbox as well as any other analytical capabilities for MSA funded research
- Attendance to all MSA Beef and Sheepmeat Pathways meetings.

2 Results

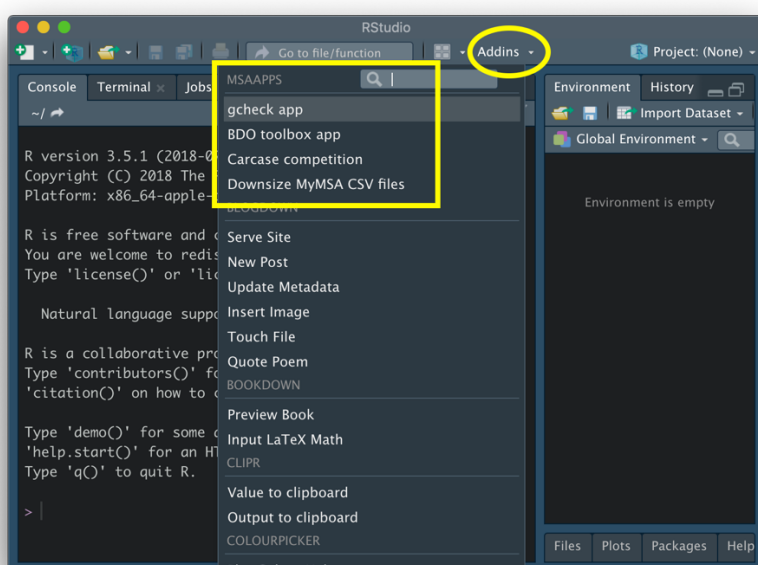
2.1 MSA staff support

Ad hoc analysis and support for:

- Jessira Perovic (BDO toolbox and ad hoc reporting)
- Will Atkinson (BDO toolbox)
- Sarah Day (BDO toolbox, caracase competition)

2.2 MSAapps R package

An R package has been developed that provides a number of apps to support the activities of Meat Standards Australia. Once installed, it provides the following dropdown menu.



2.2.1 gcheck app

The gcheck app was developed in a previous contract. It has now been incorporated directly into the MSAapps package so there is only one tool the needs to be installed and access is now easier through the dropdown menu.

2.2.2 BDO toolbox

Development of the BDO toolbox app. Implementing the fragile Excel approach in an interactive R based platform and adding new functionality. The key features are:

- Upload. Here you can upload a CSV or Excel file with the usual column names. You can also optionally filter the uploaded data set and all subsequent analyses will only use the filtered data. An overview of the (filtered) data is shown on the right with counts of the number of head in each processor and brand combination.
- Data table. To see the data that was uploaded and filtered along with any new variables that were created.
- AOR. MSA Annual outcomes report calculations.
- Non-compliance visualisation and tabular output. Daily or monthly by one categorical variable or two. Can optionally turn off the lines (and only show points) or turn off the points (to only show lines). The raw data can be extracted from the tables at the bottom.
- Traits over time
- Percentile bands
- Cattle graded by month
- Distributions
- Eating Quality by cut
- Pivot table
- Compliance heat mapping
- Member search

There is still some functionality to be implemented, e.g.

- Add in counts above and below certain (customisable) thresholds e.g. marbling, ossification and MSA index to give information that may help establish pbr lines (requested by Will last year).
- Average traits around certain MSA index quantiles (e.g. MSA marbling for animals at the 90th MSA index percentile will return the average MSA marbling for animals identified as being in the 89.5th to 90.5th percentile band).

2.3 Workshop

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. 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.

A workshop was held in May 2019, the first three days in Brisbane with only 3 attendees, the follow up two days in Armidale which had the original three attendees plus a couple of alumni from previous workshops. Feedback was positive. The full set of feedback can be found in Appendix 6.1.

2.4 Pathways attendance

Attended all pathways meetings, providing statistical advice and analysis to support discussions.

3 Conclusions/Recommendations

Recommendation 1

The BDO toolbox is reaching feature maturity, however, the unrealised potential is a direct database link that would avoid the need to download/upload with large files. This may also see the app be hosted on a server with the RAM and computing capabilities beyond any individual staff member's laptop which is often a bottleneck for analysis.

Recommendation 2

Continue with the 5 day format for initial Analytics for Industry training workshops. The dates should be finalised months in advance and advertised widely. Perhaps move back to April as the May session was not well attended. 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 3

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.

4 Appendix

4.1 Workshop feedback

	Participant 1	Participant 2	Participant 3
The presenters communicated the information clearly	Strongly Agree	Strongly Agree	Strongly Agree
The presenters made the subject matter compelling	Strongly Agree	Strongly Agree	Strongly Agree
The presenters were able to answer questions	Strongly Agree	Strongly Agree	Strongly Agree
The presenters were friendly and approachable	Strongly Agree	Strongly Agree	Strongly Agree
ABOUT THE PRESENTERS	Garth and Kevin really knew their stuff but were able to explain it in a way that was so easy to understand and apply	Everything was clear and clarified and could see the reasoning behind completing the activity and the application of it.	Garth and Kevin were great. Nothing was a problem, there were no silly questions and sidetracks were welcome. Questions were always responded to and if they didn't know it was an adventure to find the answer. They made this a really enjoyable learning experience, especially given that the content can be difficult under the best circumstances.
The workshop was relevant to me	Strongly Agree	Agree	Strongly Agree
The workshop was interesting	Strongly Agree	Strongly Agree	Strongly Agree
The workshop met my purpose in attending	Strongly Agree	Agree	Strongly Agree
I want to tell others about what was presented	Strongly Agree	Agree	Strongly Agree
The website resources were useful	Strongly Agree	Strongly Agree	Strongly Agree

The opportunity to work on my own data was useful	Strongly Agree	Strongly Agree	Strongly Agree
ABOUT THE MATERIAL	Very useful (and will be good reference material in the future)	Reference material was clear and good it can be accessed following the course.	the material was good.it was good to use data from R packages that are for learning how to manipulate and present data. It was also really good to use industry data that 'makes sense' when manipulating and presenting. In other courses that I have attended the biggest constraint has been using weird data that is difficult to make sense of. The course design has been really well thought out and can be managed for the different objective people are wanting to achieve with the course. The people in the course were wanting to achieve different things and Garth and Kevin were more than happy to facilitate this.
The duration of the workshop was appropriate	Strongly Agree	Agree	Strongly Agree
The workshop was well organised	Strongly Agree	Agree	Strongly Agree
I would recommend this workshop to others	Strongly Agree	Agree	Strongly Agree
ABOUT THE WORKSHOP IN GENERAL		The break to work on individual projects and come back and present them at the end worked well - having a break was good.	Workshop was really good. I dont think that there is ever enough time with these things. I learnt so much, but probably still need to learn much more! It was nice to have the break between the two formal sessions, however I did not get a chance to continue working on this in between, and this was the case for the other members of our group. This is not a reflection of the course but one does wonder if there would be value in doing more in the earlier days, although there are difficulties with this.
What were the best aspects of the workshop?	The first day learning the basics of R and then the ability to work with nearly one-on-one assistance on our own project	Concepts were explained in understandable terms and it was made relevant to what we would be using R for individually.	Hands on learning and running your own data
What aspects of the workshop need improvement?	N/A	NA	Nothing really, the workshop is very relevant and has benefits for persons in commercial industries and research settings.

<p>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.</p>	<p>Already have kept using it to extract information and make graphs etc for presentations, will be using it to develop breeding program advice for breeders on the value proposition of genetics</p>	<p>Removing the data work around some external facing results and this takes away the potential for a mistake and is all clear as to what needs to be done, making the data analysis more efficient.</p>	<p>Hopefully if I remember cleaning data and removing unwanted data will streamline my data management (rather than doing it the hard way in excel) and the use of ggplot to generate publication ready figures will save lots of time too. Going through the emmeans package was really good for me.</p>
<p>Do you have any suggestions about the best way to continue to support and grow your analytics skills in R?</p>	<p>Monthly catchups via web or annual bootcamps/refreshers where we just dedicate time to looking at R. E.g. where we learn about new packages, then just dedicate time to working with R with the support of the team.</p>	<p>Being able to keep in contact with questions - potentially refresher training, not 100% sure at this stage</p>	<p>All of these suggestions would be good. I think it would be great to have an annual refresher this would also then allow for improvements/modifications/new packages to be worked on and help to keep moving forward with R and hopefully prevent backwards regression into old habits with excel...</p>
<p>Any other comments?</p>	<p>NA</p>	<p>NA</p>	<p>Perhaps there could be a secondary workshop for more advances modelling for researchers if there was an appropriate level of interest</p>

4.2 MSA Apps development timeline

4.2.1 MSAapps 0.2 [2019-06-25]

- Improve non-compliance tab (including no group, i.e. overall non-compliance) and allowing download to powerpoint.
- Add in counts above and below certain (customisable) thresholds e.g. marbling, oss and msaindex to give information that may help establish PBR lines (requested by Will last year).
- Average traits around certain MSA index quantiles (e.g. MSA marbling for animals at the 90th MSA index percentile will return the average MSA marbling for animals identified as being in the 89.5th to 90.5th percentile band).
- MLA branding colours throughout.

4.2.2 MSAapps 0.1.9999 [2019-02-18]

- Updated the MLA colours so that the gold isn't used as often.
- Lines in the line plot and density plots made thicker and unfaceted histogram now fill shaded.
- Added mapproj as a formal dependency (required for the ggplot2::coord_map() function).

4.2.3 MSAapps 0.1.9999 [2019-01-24]

- Fixed number formatting on the y-axis in the distribution and cattle graded by month tabs (comma for numbers in 1000s and percentage for proportions).
- Added download to powerpoint for cattle graded by month.
- Renamed the "Feed group" as "Days on feed". There are still "No. days on feed" and a "Raw feed type" which show the original values in the data file, not the adjusted values in "Feed type" and "Days on feed".

4.2.4 MSAapps 0.1.9999 [2018-01-21]

- Edit so all dentition scores are 0, 2, 4, 6 or 8 (if they were an odd number they are now rounded up to the next even number).

4.2.5 MSAapps 0.1.9999 [2019-01-21]

- Change from reporting mean to median in the boxplot on the distribution tab.
- Added option to hide outlier dots in the boxplot on the distribution tab (they are swallowed into an extended whisker).

4.2.6 MSAapps 0.1.9999 [2019-01-10]

- Added option to display sample size and mean to boxplots in the distribution tab.

4.2.7 MSAapps 0.1.9999 [2018-01-08]

- Removed non-compliance from the list of trait options in the distribution and percentile tabs as it's a discrete variable (when not averaged over days/months) and so doesn't really fit with the type of plots/analyses that is being delivered in those two tabs.
- Added MLA colour theme(s). Most notably in the app itself and coloured the eating quality star charts as required.
- Improvements to the eating quality star chart, with adaptive titles and captions as appropriate.

4.2.8 MSAapps 0.1.999 [2018-01-08]

- Distributions tab: added boxplots, added stacked (along with the existing overlay) to density and histogram. Fixed the densities so that they are normalised by the number of carcasses in each grouping.
- Added plots to eating quality by cut.
- Added download to powerpoint buttons to the plots in the distributions tab and the newly created eating quality by cut tab. Will roll this out more broadly if it's deemed useful.

4.2.9 MSAapps 0.1 [2018-01-06]

- First major release.
- Consistency in the UI throughout.
- Global plot options for axis tick marks, vertical plot height, lines or plots (on line charts).
- Made it possible to upload Excel files (in addition to csv files).
- Interactive heat map update so that the plot doesn't reset when the settings are changed. Can now turn off the scaling of the circles proportional to the number of carcasses in each postcode. Can also adjust the lower and upper bound thresholds for colour changes - the thresholds you want to use for grain are likely to be different to the thresholds for grass.
- Cattle graded by month can now show stacked bar charts with proportions on the y-axis (i.e. each bar in each month has length 1) so can potentially see trends over time.
- Member search functionality added.