CASE 1: Cashflow of a Cow

You have to build a cash flow model of a cow, for which we have collected all the relevant data in Financial Access Dollars which is displayed up to row 29. In this country the general interest rate on savings and loans is 7%, luckily there is no inflation in the country. From row 32 onwards is where you can work.

The number of cows we want to model must be flexible (cell I9). Assume that newly bought calves are 1 year old for the first 12 months in the model. In the last month they are 5 years old, the cows will be brought to a butcher and sold for meat.

 If you find it helpful, you may assume that every month has 30 days.

Try to make the model as flexible as possible, the user should be able to change the green cells at any moment, with all the calculations adjusting automatically. Also try to use formulas and automation as much as possible.

If you wish to add lines to the calculation section, feel free to do so, you may split up the calculations of the costs and revenues as much as you like, as long as the four specified lines are included in the bottom (Income, Expenditure, Net Income and Internal Rate of Return).

Finally, you may assume that the costs are spread equally across the year.

After completing the model please write a short answer on the following questions:

1. You are advising a start-up farmer (i.e. they must buy new calves before producing milk), what would you advise from a financial perspective?
2. Would this advice change if this farmer had just received 2 calves for free? Explain your reasoning.
3. Identify 3 important risks for a farmer (even if they are not in the model). For each, also suggest a way of incorporating this risk in the model (you do not actually have to put them in the model, just explain how you would go about this and/or how you would investigate the impact in the current model).
4. How would your analysis change if once you bought a cow, it would give birth to two calves over its lifetime? You do not have to model this but explain the factors to consider and how you think each factor would impact the IRR, costs or revenues.

CASE 2: STATISTICS

This data is part of the Agricultural Survey of African Farm Households (Waha, Katharina; Zipf, Birgit; Kurukulasuriya, Pradeep; Hassan, Rashid (2016): An agricultural survey for more than 9,500 African households. figshare. https://doi.org/10.6084/m9.figshare.c.1574094), downloaded from Kaggle (https://www.kaggle.com/crawford/agricultural-survey-of-african-farm-households).

This dataset is just a subset of their actual data, where irrelevant data and some countries have been dropped. The dataset is made up of 4763 observations across 5 countries. This data gives an overview of the demographics of farmers, their income and an indicator of their health.

You are tasked with cleaning and analyzing this data. You will first work with the data, which should not take you more than an hour. Please keep your work tidy and work as efficiently as you can (i.e. less code is better).

Afterwards, please make a short Word file where you summarize your findings, show your graphs and explain them. The idea is that even though you do the analysis, you must communicate your findings to a colleague, who has to write a report for a general audience on your findings. So please be sure to explain what the graph shows, trends or interesting outliers in the data. Are there interesting differences between countries? Are there interesting outliers or is the data nicely distributed? What are the main forces driving farm income, how do they affect it and what could be the driver? The point here is not to give a very detailed explanation, the goal is to provide general information that a colleague of yours can then use to write in a report targeted at a general audience.

The data analysis should take no more than 1 hour, and the write-up about 30 minutes. We understand that you might not be able to make the perfect analysis in that amount of time. Just do whatever you can and write in the Word file what you would have liked to do/change/improve/look into if you had more time.

Please send us your Word file explaining your work and highlighting your results, as well as a copy of the file you used (if it is in R, please send us the .R file, if you use another software, please copy your work into a .txt or Word file).

This exercise will quickly take you through the three steps of data analysis.

1. Cleaning – making the data easier to analyze

- Clean up the data stored in the 8 gender columns so that (1 = Male, 2 = Female). Also change missing values to also be equal to Female (not what you should do in practice but only for this assignment).
- Relabel married1 so that 1 = Married, 2 = Never Married, 3 = Previously Married, 4 = NA.
- Create a new variable 'hh_size' which equals the number of people in the household.
- Create a new variable 'female_hh' which is 1 if at least 5 people in a household are women and 0 otherwise.

  The other variables are:

- Educ1 – years of education of the main interviewee
- Country – Burkina Faso, Egypt, Ethiopia, Ghana or Niger
- Incfarm – Total farm income in local currency
- Sickdays – the number of days the main interviewee was sick in the last year

2. Exploratory Data Analysis – see what the data is all about

- How many people were interviewed in each country?
- What is the proportion of female households in the entire dataset?
- Find a good way to visualize the distribution of farm income for Ghana only.
- Find a good way to visualize the age distribution for the entire dataset.

3. Analysis

- For Ghana: how does farm income vary by gender of the household (female_hh), their education (educ1), age (age1) and the number of days they were sick (sickdays)?