Portfolio Task 1

Sandy Pullen

2022-05-03

Portfolio Task 1

Introduction

The Tasmanian Land Records (1832-1935) dataset contains over 23,000 records relating to deeds of land grants under General Law, which is a system of land title based on common law used up to 1862 in Tasmania.

The Land records dataset, each record contains links to digital objects available at the Tasmanian Names Index held and maintained at Libraries Tasmania. The nature of these links is as follows:

1. The URL of the person in the Names Index, which then links to the next two URLS listed here,
2. The URL to the index page containing the index record,
3. The URL of the digitised image of the deed of land grant relating to that record.

The area of interest for this project is the location of Kempton, a historical town which was previously known as Green Ponds, in the southern Midlands of Tasmania.

Method

Original Dataset

The Tasmanian Land Records (1832 - 1935) dataset was downloaded in CSV format and saved with filename land.csv

Open Refine

The original dataset was opened in Open Refine and the following steps were performed.

1. Check for the location of Kempton, and verify that the earlier name of Green Ponds was not used, and that there were no misspellings for Kempton using 'text facets'.

2. Filter for Location=Kempton using 'text facets'. This reduced the dataset to 55 records.

3. Check that each column had data, using 'text facets' to check for blanks.

4. The URL fields contained combined links, so 'Edit Column > Split into several columns' was used to separate the columns into meaningful data as follows:

Original Column	Separator Character	New Columns
DIGITAL_OBJECT - URL_TEXT	|	DIGITAL_OBJECT - URL_TEXT 1 DIGITAL_OBJECT - URL_TEXT 2
DIGITAL_OBJECT - URL	|	DIGITAL_OBJECT - URL 1 DIGITAL_OBJECT - URL 2
NAME	,	LASTNAME FIRSTNAME
FIRSTNAME	<space>	FIRSTNAME1 FIRSTNAME2

Three records contained two firstnames, however, since this did not aid in identification of duplicates, the column was split using the space as a separator, and the second firstname was ignored.

The column DIGITAL_OBJECT - URL 2 contained a link to a URL on the Libraries Tasmania website which showed an image viewer, and the relevant image embedded into the viewer. For example the URL for the land grant for Thomas Croxton is:

https://stors.tas.gov.au/RD1-1-12$init=RD1-1-12P110JPG

Inspection of the html code using Firefox revealed that the direct link to the image file without the image viewer frame was:

https://stors.tas.gov.au/fetch/RD1-1-12P110JPG

So the column DIGITAL_OBJECT - URL 2 was split, using a separator of '=' to isolate the image name required (i.e. RD1-1-12P110JPG). The column was renamed 'IMAGE'.

The data file was saved as refined-land-csv-May-1.csv' in the folder for Portfolio Task 1.

Download Images from Libraries Tasmania Website

A Python script was developed to iterate (loop over) the rows of the dataset, and for each image listed, web scraping was used to retrieve the image file and save it with a file extension of .jpg to a subfolder named 'captured'. The LASTNAME and FIRSTNAME1 were added to the filename created in the python script to aid readability of the image list.

Web sources used to develop the script are noted here: [1] [2] [3]

The python script is located in the file 'capture-images.py' in the folder for Portfolio Task 1.

An example of the digitised land record image is shown in Figure 1.

Data Entry

The .csv file created in the previous step was opened in Excel and saved in Excel format as 'data-entry.xlsx'

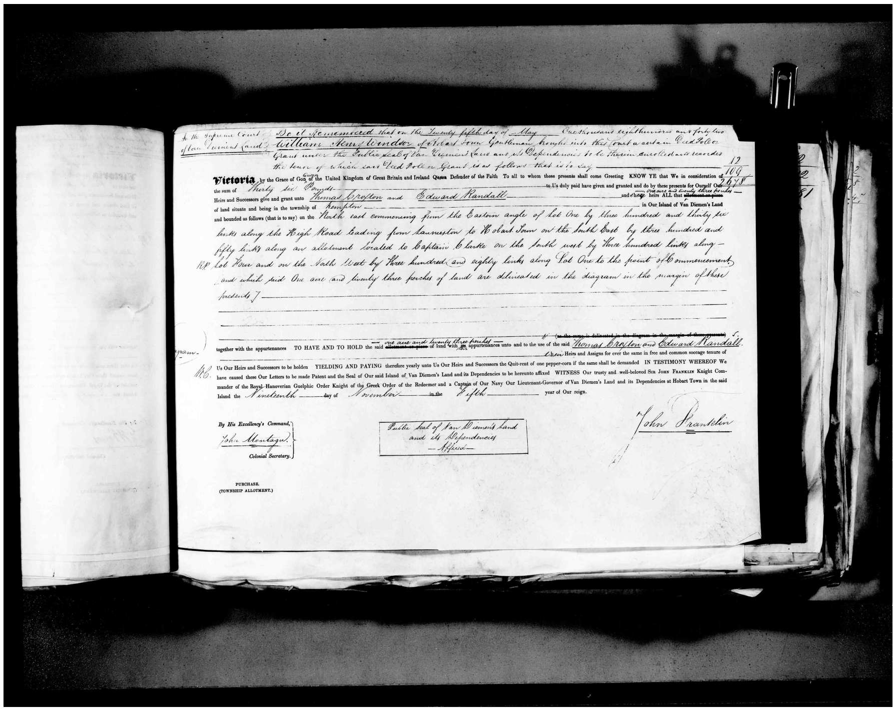
New columns were added to allow data entry of information transcribed from the digital image, as described in the data schema in ReadMe.Rmd

- SUM_POUNDS
- SUM_SHILLINGS
- SUM_PENCE

[1] https://www.rstudio.com/blog/three-ways-to-program-in-python-with-rstudio

[2] https://www.machinelearningplus.com/pandas/pandas-read_csv-completed/

[3] https://www.educative.io/edpresso/how-to-locally-save-an-image-using-urllib



In the Supreme Court of Van Diemen's Land } Be it Remembered that on the Twenty fifth day of May One thousand eight hundred and forty two William Henry Windsor of Hobart Town Gentleman brought into this Court a certain Deed Poll or Grant under the Public Seal of Van Diemen's Land and its Dependencies to be therein enrolled and recorded the tenor of which said Deed Poll or Grant is as follows that is to say — 12 / 109 / 2678

Victoria by the Grace of God Queen of the United Kingdom of Great Britain and Ireland ~~Queen~~ Defender of the Faith To all to whom these presents shall come Greeting KNOW YE that We in consideration of the sum of Thirty six Pounds — to Us duly paid have given and granted and do by these presents for Ourself Our Heirs and Successors give and grant unto Thomas Croxton and Edward Randall — and their heirs ALL that ~~allotment or piece~~ One acre and twenty three perches of land situate and being in the township of Kempton — in Our Island of Van Diemen's Land and bounded as follows (that is to say) on the North east commencing from the Eastern angle of Lot One by three hundred and thirty six links along the High Road leading from Launceston to Hobart Town on the South East by three hundred and fifty links along an allotment located to Captain Clarke on the South west by three hundred links along Lot Four and on the North West by three hundred and eighty links along Lot One to the point of Commencement and which said One acre and twenty three perches of land are delineated in the diagram in the margin of these presents] —

~~(as the same is delineated in the diagram in the margin of these presents)~~ together with the appurtenances TO HAVE AND TO HOLD the said ~~allotment or piece~~ one acre and twenty three perches of land with ~~its~~ their appurtenances unto and to the use of the said Thomas Croxton and Edward Randall their Heirs and Assigns for ever the same in free and common soccage tenure of Us Our Heirs and Successors to be holden YIELDING AND PAYING therefore yearly unto Us Our Heirs and Successors the Quit-rent of one pepper-corn if the same shall be demanded IN TESTIMONY WHEREOF We have caused these Our Letters to be made Patent and the Seal of Our said Island of Van Diemen's Land and its Dependencies to be hereunto affixed WITNESS Our trusty and well-beloved Sir John Franklin Knight Commander of the Royal Hanoverian Guelphic Order Knight of the Greek Order of the Redeemer and a Captain of Our Navy Our Lieutenant-Governor of Van Diemen's Land and its Dependencies at Hobart Town in the said Island the Nineteenth day of November in the Fifth year of Our reign.

By His Excellency's Command,
John Montagu
Colonial Secretary.

Public Seal of Van Diemen's Land
and its Dependencies
— Affixed —

John Franklin

PURCHASE.
(TOWNSHIP ALLOTMENT.)

Figure 1: Figure 1: Deed of Land Grant for Thomas Croxton - image RD1-1-12P110JPG

- BOUNDARY DESCRIPTION
- DIAGRAM

Data Validation

Data Validation was enabled for the DIAGRAM column to allow values of 'yes' or 'no'.

Results

For the 55 records, there were 54 unique images. One image was repeated due to the land being owned in partnership. (i,e, Thomas Croxton and Edward Randall were joint owners, so there is one record for each person, linking to the same image.)

One record had blank year information, this was visually checked on the deed image, and confirmed as missing. 'missing' was entered into this cell in the data entry phase.

Not all records

Reference List