Earnest Analytics Code Test

We have included CSV data in the `back_enriched_addresses.csv` file attached. Please spend approximately 2 hours on the test.

Any further questions, you can contact Nick Rogers on nrogers@earnestanalytics.com.

Write a program or script that can perform data transformations on an input file. We like to see tests where possible. If you can't complete a step that's fine, move on to the next. The steps in bold are stretch goals for extra credit - feel free to skip them if you don't have the time. Reorder steps if it makes more sense to you to do so.

1. You can use the attached CSV as a sample input that your program or script should be able to process
2. Perform as many of the following transformations as you can:
    a. Remove all rows where pdays is -1
    b. Split name into first name and second name columns (drop name)
    c. Replace the values in the age column with bucketed values, such that x < 10 becomes 0, 10 <= x < 20 becomes 1, etc.
    d. Replace yes/no values with booleans
    e. Replace day and month with a single date column, of the form dd/MM
    f. Rename the y column to "outcome"
3. **Add a column which categorizes geographical features in the address, where present. Note the dirtiness of the address data and that the exact categories:**
    a. "water", where the address contains e.g. lake, creek
    b. "relief", where the address contains e.g. hill, canyon
    c. "flat", where the address contains e.g. plain
4. Group by the feature (if you created it, or by some other field if not) and filter out any empty values, sort by the age bucket (or age if you didn't do the bucketing), and return a row count.
5. Write the row level data from step 2, and aggregated data from step 3 to both CSV and parquet formats.

Please document the following in a README.md file:

1. Instructions for reviewers to build and test your work.
2. Implementation and design choices you made.
3. Any further work you would like to do to make it production ready.