

1. *Create a ***tidy*** spreadsheet/table listing the names of Danish monarchs with their birth- and death-date and start and end of their reign. They should be sortable by year of birth. Suitable source website is for example [here](#), but you can also use another source, provided you reference it.*

I have attached my spreadsheet in the submission.

I used <https://www.kongehuset.dk/monarkiet-i-danmark/kongerakken> and <https://danmarkshistorien.dk/vis/materiale/kongeraekken> as my source.

2. *Does OpenRefine alter the raw data during sorting and filtering?*

No, it does not. OpenRefine changes the way we visualize the data, but it does not alter the raw data in any way.

Even if we for example create a new facet and use the Value.replace command, then we are only highlighting and changing the visual look of the data, but the raw data remains the same.

- 3 *Fix the [interviews dataset](#) in OpenRefine enough to answer this question:
"Which two months are reported as the most water-deprived/driest by the interviewed farmer households?"*

1. **First, I made a Facet that would show every month, however as you can see on the picture it is out of order because of all the symbols.**

OpenRefine SAFI openrefine (1) csv [Permalink](#)

Facet / Filter Undo / Redo 1 / 1 **131 rows**

Refresh Reset all Remove all

months_no_water change

17 choices Sort by: name count Cluster

[Apr'; 'May'; 'June'; 'July'; 'Aug'; 'Sept'; 'Oct'; 'Nov'] 1
 [Aug'; 'Sept'; 'Oct'; 'Nov'; 'Dec'] 4
 [Aug'; 'Sept'; 'Oct'; 'Nov'] 10
 [Aug'; 'Sept'; 'Oct'] 11
 [Aug'; 'Sept'] 6
 [Jan'; 'Dec'] 2
 [July'; 'Aug'; 'Sept'; 'Oct'; 'Nov'; 'Dec'] 1
 [Nov'; 'Dec'] 2
 [Nov'] 1

1. Facet Text facet
 Text filter Numeric facet
 Edit cells Timeline facet
 Edit column Scatterplot facet...
 Transpose Custom text facet...
 Sort... Custom numeric facet...
 View Customized facets
 Reconcile 17-Nov-16 5

2. NULL 17-Nov-16 6

2. Then I went into Edit cells -> Transform

months_no_water interview_date quest_no

Facet 17-Nov-16 1

Text filter

Edit cells Transform...

Edit column Common transforms

Transpose Fill down

Sort... Blank down

View Split multi-valued cells...

Reconcile Join multi-valued cells...

Cluster and edit...

Replace...

NULL

3. In this menu I entered the expression:

`value.replace("[", "").replace("]", "").replace("'", "").replace(" ", "")`

Custom text transform on column months_no_water

Expression Language General Refine Expression Language (GREL)

```
value.replace("[", "").replace("]", "").replace("'", "").replace(" ", "")
```

No syntax error.

Preview

History

Starred

Help

row	value	value.replace("[", "").replace(...
1.	NULL	NULL
2.	[Aug' ; 'Sept]	Aug;Sept
3.	NULL	NULL
4.	NULL	NULL
5.	NULL	NULL
6.	NULL	NULL

On error ☒ keep original ☐ Re-transform up to times until no change
☐ set to blank
☐ store error

OK Cancel

4. This removes every special character and makes my new facet look like this:

months_no_water

change

17 choices Sort by: name count Cluster

- Apr;May;June;July;Aug;Sept;Oct;Nov 1
- Aug;Sept 6
- Aug;Sept;Oct 11
- Aug;Sept;Oct;Nov 10
- Aug;Sept;Oct;Nov;Dec 4
- Jan;Dec 2
- July;Aug;Sept;Oct;Nov;Dec 1
- Nov 1
- Nov;Dec 2
- NULL 45
- Oct 2

- However, it is still not tidy.

5. To change this I went into Facet -> Custom text facet.

- Then I entered the command `value.split(";")`

months_no_water interview_date ques

Facet ▶ Text facet

Text filter Numeric facet

Edit cells ▶ Timeline facet

Edit column ▶ Scatterplot facet...

Transpose ▶ Custom text facet...

Sort... Custom numeric facet...

View ▶ Customized facets

Reconcile ▶

Custom facet on column months_no_water

Expression Language General Refine Expression Language (GREL) ▼

value.split(";") No syntax error.

Preview History Starred Help

row	value	value.split(";")
1.	NULL	["NULL"]
2.	Aug;Sept	["Aug", "Sept"]
3.	NULL	["NULL"]
4.	NULL	["NULL"]
5.	NULL	["NULL"]
6.	NULL	["NULL"]

OK Cancel

- This removes the ; from the facet and finally makes the months sortable.

6. Lastly, I can now see which month is the driest by sorting by count.



The screenshot shows a window titled 'months_no_water' with a 'change' button. Below the title bar, it says '11 choices Sort by: name count'. The table lists months and their corresponding counts, sorted in descending order. The months are: Oct (74), Sept (70), Nov (51), NULL (45), Aug (33), Dec (11), Jan (2), July (2), Apr (1), June (1), and May (1). At the bottom, there is a facet bar labeled 'Facet by: choice counts'.

Month	Count
Oct	74
Sept	70
Nov	51
NULL	45
Aug	33
Dec	11
Jan	2
July	2
Apr	1
June	1
May	1

- October and September are the two driest months.