

WEX Bulk Student Report in Argos: Tidying Template

AUTHOR

Andrea Puglisi

PUBLISHED

August 26, 2025

Context

Working with university IT, IT developed **WEX Bulk Student Report** in Argos, the university's new reporting system, which adheres to our bulk patron template required by OCLC. In addition, university IT was able to create a template to pre-fill columns such as borrowerCategory (Student) and idAtSource, while removing dateOfBirth, gender, address, mobile phone, information automatically. This is a significant improvement.

Script Summary

The script below can and does a few required tidying tasks, such as ensuring that nickname cells only contain contents which differ from givenName and setting the oclcExpirationDate. The oclcExpirationDate must be manually set for 07-01, 5 years into the future, following YYYY-MM-DD format.

About the Process of Importing and Exporting Bulk Student Data into RStudio

Instructions on how to download the report from Argos are included on the library's Systems Intranet Guide.

The WEX Bulk Student Report downloads as a .csv file (no.xlsx option is available). It is recommended to open the file in Excel before importing into R for this process. Please save the .csv report as an .xlsx *before* completing this process.

The R script writes two .xlsx files. Following the completion of this process, open the tidied file in Excel and make sure it looks as expected (check dates, and nicknames). Then, save the file as a tab-delimited text file onto computer, and after saving, change the file extension to a .tsv per OCLC. This is the same step that must be completed if using the process to tidy in Excel.

Template Instructions: Follow the #Comments & update file names

```
library(openxlsx)
library(tidyverse)
```

— Attaching core tidyverse packages — tidyverse 2.0.0 —

```
✓ dplyr      1.1.4    ✓ readr      2.1.5
✓ forcats    1.0.0    ✓ stringr    1.5.1
✓ ggplot2    3.5.1    ✓ tibble     3.2.1
✓ lubridate  1.9.4    ✓ tidyr      1.3.1
✓ purrr      1.0.4
```

— Conflicts — tidyverse_conflicts() —

```
* dplyr::filter() masks stats::filter()
* dplyr::lag()     masks stats::lag()
```

i Use the conflicted package (<<http://conflicted.r-lib.org/>>) to force all conflicts to become errors

```
library(readxl)
library(dplyr)
```

#Step 1: Read in the data

```
fall_2025_student_list <- read_excel("Student Patrons WEX Test2 Report_20250825_115315_2x.xlsx") %>%
  #Update file names
  mutate(across(everything(), as.character)) %>%
  mutate(across(everything(), ~replace(.x, is.na(.x), "")))
```

New names:

- `` -> `...7`

```
View(fall_2025_student_list) #Update file name to match above

#Step 2, PreferredNames Report, helpful for Head of Access Services/UX Librarian:
preferred_names_fall_2025_list <- fall_2025_student_list %>% #Update file name to match Step 1
  structure
  filter(givenName != nickname)
  View(preferred_names_fall_2025_list) #Update filename

#Step 3, Save Preferred Names Report
write.csv(preferred_names_fall_2025_list, "2-2025_fall_preferred_names_list.csv") #Update filenames

#Step 4, Ensures that prefilled nickname cells only contain contents that differ from givenName;
  Manually set oclcExpirationDate to 4 years into the future

fall_2025_students_nicknames_added <- fall_2025_student_list %>% #Update filenames
  mutate(nickname = if_else(nickname != "" & nickname != givenName, nickname, ""),
    oclcExpirationDate = as.Date("2029-07-01"), #Set OCLC expiration date manually
    oclcExpirationDate = format(oclcExpirationDate, "%Y-%m-%d")) %>% #Ensure format is YYYY-MM-DD as
    character when you View sheet
  mutate(across(everything(), as.character))

#Review, export, and then open in Excel before converting to tab-delimited .tsv for OCLC
View(fall_2025_students_nicknames_added) #Update filename
write.xlsx(fall_2025_students_nicknames_added, "wex_2025_fall_students.xlsx", rowNames = FALSE)
  #Update filenames
```

Do Not Edit: Sample Template

This is a back-up copy of R script that can be used to identify patron's preferred names from Argos Bulk Student Report, and set patron account expiration date. This template uses Fall 2025 as an example and contains detailed explanatory information about the script.

RStudio Set-up Requirements

Create a patron uploads folder, and a Bulk Patron Upload RStudio Project.

When doing bulk patron uploads, open the Bulk Patron Upload project so that settings and installed packages load. The Bulk Patron Upload Project should have the tidyverse, readxl, writeexcel, dplyr packages installed.

```
library(openxlsx) #needed to write XSLX
library(tidyverse) #Download via CRAN and install in project environment
library(readxl) #Need this package in order to "read in" .xlsx files
library(dplyr) #Need this in order to use the pipe, relocate, mutate and filter functions

#Step 1: Read in the data; remove NAs that populate upon import

fall_2025_student_list <- read_excel("Student Patrons WEX Test2 Report_20250825_115315_2x.xlsx") %>%
  #This imports the file into RStudio; the %>% symbols form what is called a pipe, and
  effectively says "take this, and then do that" -- it is a connecting operator
  mutate(across(everything(), as.character)) %>% # convert all columns to character
  mutate(across(everything(), ~replace(.x, is.na(.x), ""))) # replace any NA with blank
  View(fall_2025_student_list) #View the file in its own window in RStudio. Good idea to look at it
  and make sure it looks okay.

#Step 2, PreferredNames Report, helpful for Head of Access Services/UX Librarian: This creates a
  report that contains preferredNames by using the filter() function from library(dplyr).
preferred_names_fall_2025_list <- fall_2025_student_list %>%
```

```
filter(givenName != nickname)
View(preferred_names_fall_2025_list) #Using View() in console to show results in RStudio as its own
document
```

#Step 3, Save Preferred Names List:

```
write.xlsx(preferred_names_fall_2025_list, "2025_fall_preferred_names_list.xlsx") #Save a .csv of
preferred names to your computer
```

#Step 4, In this iteration, IT was able to create a report in Argos that prefills "nickname" column with Preferred Names. In the university system, Preferred Names will maintain the student's "givenName" as their Preferred Name unless specified. Based on how the "nickname" column contains Preferred Names (including givenName and alternate), it is best to remove redundant data, so that only names that differ from givenName appear remain in the nickname field. The following step attempts to accomplish this task!

```
fall_2025_students_nicknames_added <- fall_2025_student_list %>%
  mutate(nickname = if_else(nickname != "" & nickname != givenName, nickname, ""), # #Only keep
  nicknames that differ from givenName
  oclcExpirationDate = as.Date("2029-07-01"), #Set OCLC expiration date manually
  oclcExpirationDate = format(oclcExpirationDate, "%Y-%m-%d")) %>% #Ensure format is YYYY-MM-DD as
  character
  mutate(across(everything(), as.character)) #keep all columns as characters. This is needed in order
  to maintain earlier direction to convert into characters; ideal for keeping numerics
  registering as characters and compatible with OCLC requirements.
```

#Review, export, and then open in Excel before converting to tab-delimited .tsv for OCLC

```
View(fall_2025_students_nicknames_added) #Preview in RStudio
```

```
write.xlsx(fall_2025_students_nicknames_added, "ap-test_wex_2025_fall_students.xlsx", rowNames =
FALSE) #Export and saved on computer
```