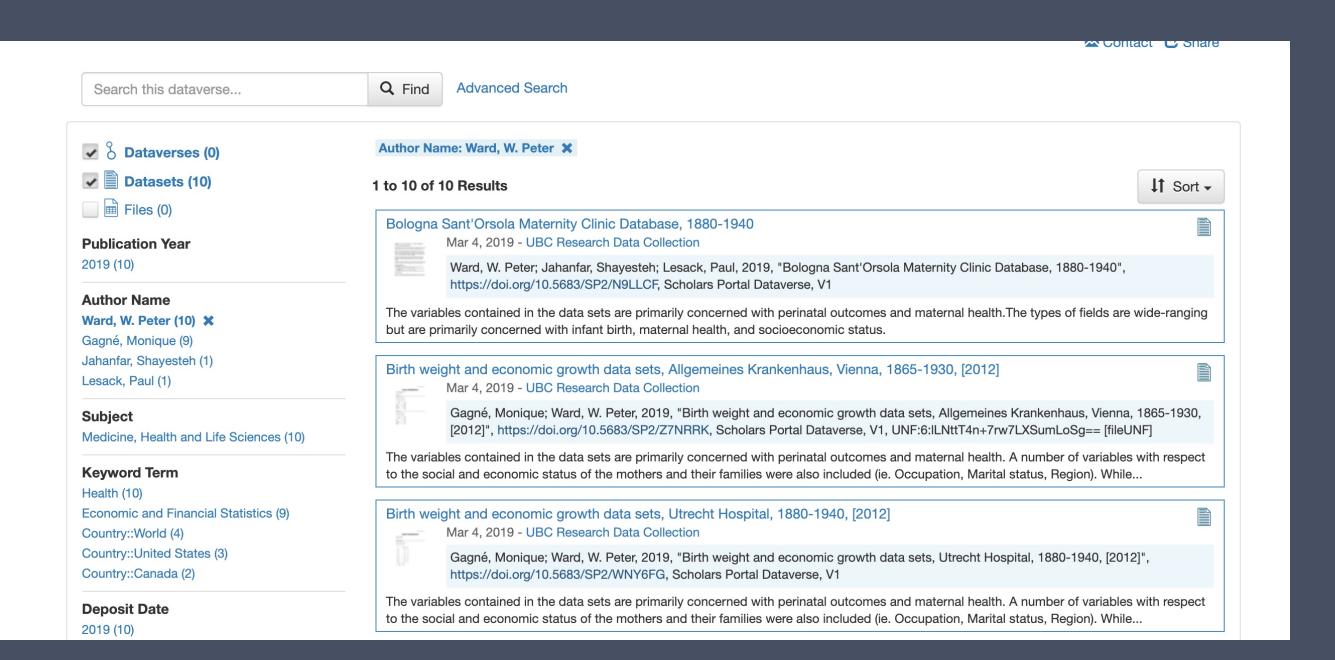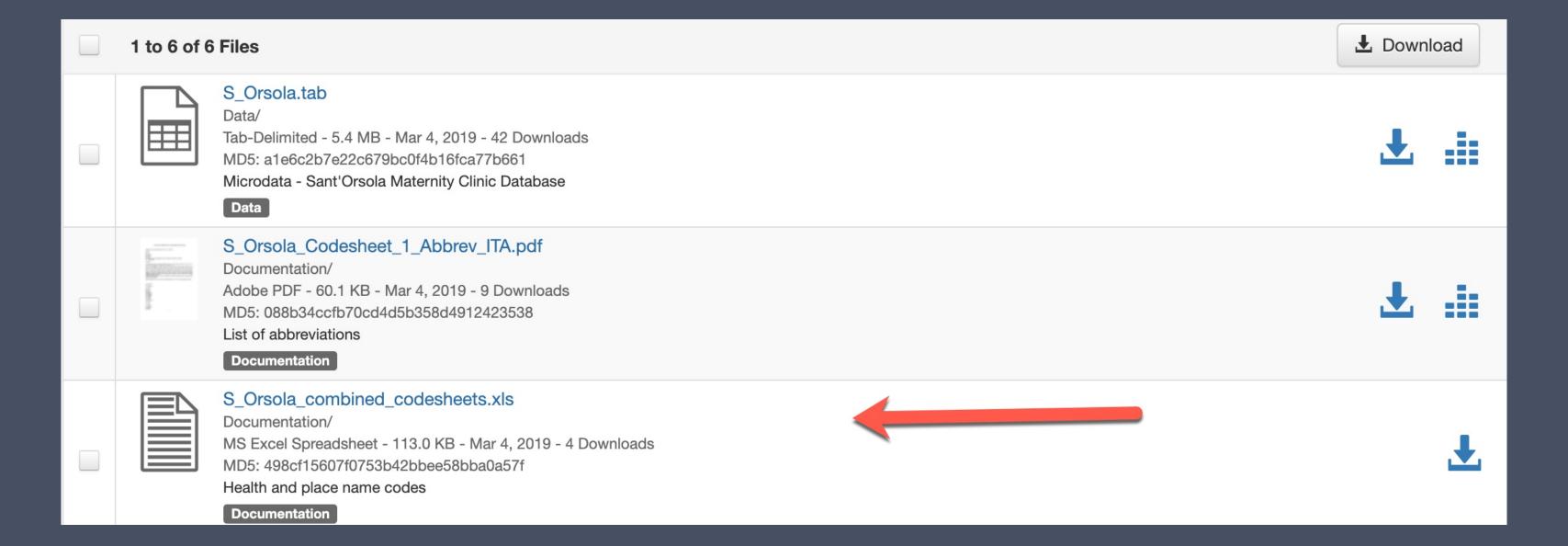DATAVERSE @UBC LIBRARY

OVERVIEW

> RESEARCH DATA - SP DATAVERSE

> MEDIATED DEPOSITS - CURATION - ALLOCATION OF DATAVERSES

> LICENSED DATA - ABACUS DATAVERSE

> AUTHENTICATION - MANAGING DATA - API USE

> DRYAD2DATAVERSE PLATFORM

> HARMONIZING DATA HOLDINGS

Research Data

> UBC Library started to work on RDM in 2014
>> First deposits were **FULLY** curated

RESEARCH DATA

RESEARCH DATA

> HIRED A PHD STUDENT TO DO DATA CURATION

> TOOK A LOT OF TIME AND RESOURCES, DATA DICTIONARIES, RECODING

> DATA MANAGEMENT VS. DATA CURATION

RESEARCH DATA

RESEARCH DATA

> FAST FORWARD SIX YEARS

> MEDIATED DEPOSIT

> USERS CAN LOGIN AND DEPOSIT DATA

> RDM LIBRARIAN / ANALYST REVIEW DEPOSITS AND CONSULT WITH USERS TO IMPROVE

> DATA IS PUBLISHED

RESEARCH DATA

> CURATION

> EVERY DATASET NEEDS LOVE

> WE HAVEN'T SEEN ONE THAT DOES NOT

> COMMONS ISSUES: LACK OF DATA DICTIONARIES – CORRUPT DATA FILES – LACKING METADATA – NO GEOSPATIAL METADATA

RESEARCH DATA

DATA CURATION, LEVEL 2:

ACTIVITIES THAT ENHANCE THE DISCOVERABILITY OF DATASETS AND HELP ENSURE THEIR USABILITY OVER TIME. E.G., RECOMMENDED METADATA FIELDS ARE POPULATED AND DATASET INCLUDES SUFFICIENT DOCUMENTATION TO ALLOW A USER WITH A SIMILAR BACKGROUND TO UNDERSTAND THE DATASET AND OPEN AND USE THE FILES.

RESEARCH DATA

Hi

I apologize for the delay in the publication of your material,

I had a look at the study which you submitted and I have some suggestions which I hope you will consider.

1. A data dictionary for your CSV file outlining what each variable represents would be very helpful for researchers using your data. Normally this information is included in a file traditionally called README. More information on files of this type is available here:

https://researchdata-06oct2014.sites.olt.ubc.ca/files/2020/04/QuickGuide_UBC_readme_v1.0_20200427.pdf

2 As well, your data set is listed as a CSV, but the separator for values is a semi-colon, which makes it not really a csv file (ie, comma-separated value). As a result, the automated conversion to a .tab file (really just a TSV) has failed because the processor is expecting a comma as a separator. You can see the result of this by viewing your data set and selecting "Explore". You will see that it doesn't load correctly as tabular data. You should consider recreating the file as a csv or tsv (tab-separated) and re-uploading it.

RESEARCH DATA

> WE DO ALLOW (AND ENCOURAGE) SOME RESEARCH GROUPS TO CREATE AND MANAGE THEIR OWN DATAVERSES

> EXAMPLES:

> HOUSING RESEARCH COLLABORATIVE (20 DATASETS)

> CENTRE FOR SUSTAINABLE FOOD SYSTEMS (22 DATASETS)

> DR. PAUL PAVLIDIS DATAVERSE (36 DATASETS)

> MANY MORE...

LICENSED DATA

DRYAD2DATAVERSE

> AT ITS SIMPLEST, DRYAD2DATAVERSE IS A STANDALONE PROGRAM THAT AUTOMATICALLY HARMONIZES RESEARCH DATA BETWEEN DRYAD TO DATAVERSE.

> REQUIRES NO EFFORT FROM RESEARCHERS WHO HAVE DEPOSITED INFORMATION INTO DRYAD, MAKING THE INSTITUTIONAL PARTNERSHIP MORE VALUABLE.

DRYAD2DATAVERSE

> HATE CODING? -- DRYAD2DATAVERSE ALSO COMES WITH A COMMAND LINE TOOL, NO KNOWLEDGE OF PYTHON OR PROGRAMMING IS REQUIRED

> GET STARTED

> `pip install git+https://github.com/ubc-library-rc/dryad2dataverse.git@master`

DRYAD2DATAVERSE

> [CODE](#)

> [DOCUMENTATION](#)

QUESTIONS?