1. ETL project files copied to Project2 folder.
2. XL files changed (and moved out of data folders to Project2 folder)
   1. Glossary – expanded
   2. Facilities – manually filled in any missing location data.
3. Evaluated other possible fields of interest in raw data.
   1. chose pl\_insol (Insolation Flux [Earth Flux]).
   2. Added to column\_descriptions and new\_column\_names XL files.
   3. Changed ETL-Project code to include.
   4. Needed to download new chrome driver for chrome86.
   5. Ran ETL pandas pgm, loaded postgres.
   6. Also added habitability code to planets, set all to 0. Will be filled in with a 1 or 2 from later input file of potentially habitable planets.
4. Manually created potHabitPlanets.xls from two tables on <http://phl.upr.edu/projects/habitable-exoplanets-catalog/> . planets from first table given habit\_code of 1, “more likely to have a rocky composition and maintain surface liquid water (*i.e.* 0.5 < Planet Radius ≤ 1.5 Earth radii or 0.1 < Planet Minimum Mass ≤ 5 Earth masses)”. Planets from second table given habit\_code 2, “less likely to have a rocky composition or maintain surface liquid water (*i.e.* 1.5 < Planet Radius ≤ 2.5 Earth radii or 5 < Planet Minimum Mass ≤ 10 Earth masses)”.

Tested the list in modified ETL/main.ipynb by merging with raw data file. Corrected 5 records to match raw data. Good run, new file has 60 records, all matched, no NaNs.

1. Added code to main.ipynb to outer join potHabitPlanets with main\_df (extracts from raw input), deleting 2nd “planet” column and converting NaNs in the habit\_code to 0 and set column to integer.
2. Added habit\_code and solar\_flux to table\_definitions.sql via quickdatabasediagrams and created new ERD image file. Ran table\_definitions.sql in postgress planetary db. Ran main.ipynb in entirety, 07 Load Postgres failed for stars table dependent keys, moved to after planets and ran fine. Replaced glossary.xlsx in with updated version from Project2 folder, reran main.ipynb top to bottom no issue.
3. Found and corrected error in centigrade/Fahrenheit calculations from kelvin with regard to planetary equilibrium temperature in main.ipynb/04 Table Split-Planet.

10/12/20 ETL-Project folder

1. code updated with changes for new fields.
2. postgres sql and ERD updated.
3. Project2/Facilities.xlsx has been manually finished, needs load to postgres table.
4. Project2/glossary will also need to be reloaded to postgres following any updates.

**Galcoordinate chart**

* D3-Project code modified to chart galactic coordinates as axes and show exoplanets.
* Also used stars as input and also showed both

1. Would like to modify code to give user choice of displays; stars, all exoplanets, potentially habitable planets (break down by code 1 and 2).
2. On hover display name, distance, galcoordinates, (size, density(relative to earth or sun), temp?, solar flux,???)

**Sunburst chart?**

* Recreate the first chart from PHL showing planet distances from stars based on insolation flux received but use sunburst graph instead of cartesian chart with background image showing ranges.

10/22 am

* Downloaded **phl\_exoplanet\_catalog.csv** from <https://github.com/jeffhoffmanmba/inhabitable_exoplanets_grp_prjt_2.git> and added to data\_in folder in ETL-Project folder
* Manually deleted all columns save name, P\_HABZONE\_OPT, P\_HABZONE\_CON, and P\_TYPE\_TEMP.
* Created single P\_HABZONE column from \_OPT column, changing anything with a ‘1’ in \_CON column to a value of ‘2’.
* Saved as phl\_habzone.csv.
* In main.ipynb

1. created 03.M2 section. Merged phl\_habzone.csv with main dataframe, adding columns 29 and 30 (habzone and temp). (verified data in xl, keeping null values for planets not on phl table for new columns).
2. Added new columns to planet\_df, renaming as habzone and temp\_class

* lsdh