

Sloan Digital Sky Survey Software Development: Logbook

Riley Thai, Andy Casey (Supervisor)

February 14, 2024

Added MWM dummy file tests and mask fixes (Monday, Feb 12, 2024 - 11:30)

I added MWM dummy file tests and fixed mask issues, which was made from Andy's code. This includes forcing masks to be boolean prior to entering the Spectrum1D initializer, and adding checks to see if the file is empty for MWM files based on DATASUM in the HDU header.

Alongside this, I resolved all the main conflicts, so that it's ready for review by rosteen this week!

Potential Bug in Documentation (Thursday, Feb 8, 2024 - 21:31)

Documented a potential bug in the plotting code. Also removed some unused presentation material.

Work on Week 9 Progress Report (Tuesday, Jan 30, 2024 - 18:54)

Made progress on the week 9 progress report, implementing fixes.

Fix: Remove 'trace0' from Hoverinfo (Tuesday, Jan 30, 2024 - 09:43)

Fixed an issue by removing 'trace0' from hoverinfo. Achieved by setting the name to an empty string on scatter traces. Also removed this task from the todo list.

Add: Resets on Expression + Flags, Colorbar Label in Aggregated (Tuesday, Jan 30, 2024 - 09:41)

Implemented resets on expression and flags, as well as improvements to the colorbar label in aggregated plots so one doesn't need to open the settings to see what is being colored (mean, min max?, etc). The resets now occur when the loaded dataset is changed. Note that the flag feature might be removed in the future, since it's should probably be kept across different dataset changes.

Fix: Delete All Views on Dataset Change (Tuesday, Jan 30, 2024 - 09:05)

Fixed a bug where all views were not being deleted upon dataset change. Now, all views are properly deleted. This is a temporary fix while exploring a more fluid dataset change protocol.

Fix: Remove Context Menu (Monday, Jan 29, 2024 - 17:07)

I had to remove this context menu in the skyplot causing performance slowdowns. The grid context might be the cause, but further investigation is needed.

Add: Memo'd Binning Operations (Monday, Jan 29, 2024 - 11:00)

Implemented memoization for binning operations to optimize performance. This prevents expensive computations when non-data parameters are modified, such as colorscale, flips, and logarithmic axes. It will save the currently binned data outputs, and use them unless explicit data properties are changed.

Work on State Lag Fixes (Sunday, Jan 28, 2024 - 11:51)

Worked on fixing state lag issues. Moved plot settings and state directly into functions to optimize performance. No longer using instances of classes; instead, everything is programmed directly.

Feat: Dynamic Subset Rerender on Skyplot (Sunday, Jan 28, 2024 - 11:28)

Implemented dynamic subset rerendering on skyplot to enhance user experience. Currently experiencing lags, aiming to resolve them.

Add: Favicon (Thursday, Jan 25, 2024 - 11:02)

Added a favicon, although it's currently non-functional. Monitoring a discussion in the Solara Discord for further guidance.

Work: Week 9 Progress Report Complete (Thursday, Jan 25, 2024 - 10:43)

Completed the progress report for week 9.

Add: Week 9 Report (Wednesday, Jan 24, 2024 - 22:19)

Started working on the week 9 report.

Fix: Dark Theme Colors, Card Height Lock, Filter Unset on Derender (Wednesday, Jan 24, 2024 - 11:28)

Added the to dark theme colors to each plot, so they match nicely against the dark theme now. I also made the card height locked, so it can only be resized horizontally. This makes sense, since it doesn't properly resize anyways vertically.

I also added a callback to force the filter to unset on derender. The application's is still slow, probably something to do with the multiple datasets, but I did fix a few things with it.

Update: Readme Specutils PR Link (Tuesday, Jan 23, 2024 - 20:52)

Updated the readme to fix the link related to specutils pull request.

Feat: Dark Theme Init, Multiple Datasets (Tuesday, Jan 23, 2024 - 15:04)

Implemented dark theme via solara's run command (`-theme-variant=dark`), and added the ASPCAP and the Cannon datasets, with options in the topbar to change them.

I also started to work on the adaptive zoom in `sbug.py` – currently trying to figure out how to loop it specifically around -180 to 180 deg correctly.

Fix: Stride Bug Bypass (Monday, Jan 22, 2024 - 12:58)

Resolved a stride bug by implementing a bypass. Because the minmax routine seems to be bugged when filters are applied, we manually take the minimum and maximum values directly and converting them to floats.

This is, computationally, horrible, but hopefully we can find out why it breaks at some point.

Feat: Histogram Selections (Monday, Jan 22, 2024 - 12:23)

Implemented histogram selections, allowing users to interactively select bins. Default bin count adjusted to 200 for improved usability.

Fix: Skyplot Relayout + Margins (Thursday, Jan 18, 2024 - 12:24)

Fixed issues related to skyplot relayout and margins, ensuring proper alignment and behavior.

Corrected delete functionality on gridlayout (Thursday, Jan 18, 2024 - 12:24)

I implemented a correction to the delete button on each card in the GridLayout, ensuring that items are correctly deleted from each render. However, in memory, there are still a bunch of figure widgets, although they don't consume much memory since they're not in the render. Nonetheless, I should change this in future

Added delete button to table and removed print debugs (Wednesday, Jan 17, 2024 - 12:38)

I added a delete button to the table ViewCard and removed the print debugs across the gridlayout.

Adjusted quick filter flag to placeholders (Wednesday, Jan 17, 2024 - 11:58)

I adjusted the quick filter flags to use placeholder flags.

Currently, it applies nothing to this dataset due to flags with meanings I don't know. I need to find these meanings on the wiki and ensure I filter out only incorrect or inaccurate data.

Implemented quick filter menu (Tuesday, Jan 16, 2024 - 11:44)

I implemented a quick filter menu with options for signal-to-noise ratio greater than 50 and drop flag. However, the flag regex doesn't detect all columns as expected – the regex `vaex` uses internally is not standard regex – it seems to not have the global flag, and I can't add it in either.

Enabled astra file read from local directory (Tuesday, Jan 16, 2024 - 10:54)

I enabled reading Astra files from the local directory. It could also be read via SQL from somewhere else, but that's a later concern.

Updated README with basic descriptions (Tuesday, Jan 16, 2024 - 10:50)

I updated the README with basic descriptions.

Added griddraggabletoolbar without modification on Solara (Tuesday, Jan 16, 2024 - 10:46)

I added `griddraggabletoolbar` separately using identical source code without requiring modification of the Solara source to run/create this. `component_vue` did not work, so I had to do it by literally copying it and just changing a line in the vue source.

I'm becoming more concerned about the number of imports in this file and whether it will slow down anything.

Partially fixed dynamic relay layout and render bug with flip/log on scatter (Monday, Jan 15, 2024 - 13:09)

I partially fixed a bug related to dynamic relay layout and render with flip/log on scatter plots. More work is needed on this, specifically because the fix doesn't really work, only for logarithmic axes *sometimes* does it work.

Added run.sh files for running with bash (Monday, Jan 15, 2024 - 11:58)

I added `run.sh` files for running with bash. Additionally, I moved `todo.md` to the main directory.

Added datatable with sort viewcard to gridlayout (Sunday, Jan 14, 2024 - 12:23)

I added a datatable with ascending and descending sorting options as a `ViewCard` to the grid. This addition enhances the user experience, since they can sort their view of the dataset.

Implemented expreditor 3 part expressions and removed brackets requirement for AND expressions (Sunday, Jan 14, 2024 - 11:35)

I implemented expreditor 3 part expressions, which now work as expected.

Additionally, I removed the requirement for brackets for AND expressions, which improves the user experience. I also made a note to consider adding unit tests for this feature. Furthermore, I fixed a bug in the `skyplot`. I removed some AST testing and AST import during the commit squash.

On the 3part, the code is really horrible, using so many `strip` and `replace` calls. I'm hoping to revise the entire expression editor thread process in future to not be so tedious, but still having validity.

I hate Jupyter (Friday, Jan 12, 2024 - 12:07)

After spending considerable time setting up the development environment and making changes to the `resize` event observer in `TypeScript`, it did not affect the rendering as expected. I am angry.

Implemented draggable panels via toolbar (Thursday, Jan 11, 2024 - 13:37)

To make the panels draggable, I added a toolbar to the top. This enhancement required adding a v-toolbar to the Vue specification. But I did this by modifying the Solara source, instead of pulling it from a component_vue or other source. The problem is that this forcefully offsets the corner resize drag bit, but I think I can fix this in the vue spec as well.

Fixed panels to menu transition (Wednesday, Jan 10, 2024 - 12:27)

I fixed the transition from panels to menu, preventing the slide down the expansion panels used to do. However, there are still other fixes that need to be addressed.

Implemented resizable and draggable grid layout (Wednesday, Jan 10, 2024 - 11:18)

I implemented a resizable and draggable grid layout. This actually amazing.

All plotting views are panels, arranged on an invisible grid, and are now added via a menu, and each panel has its own individual options, which means a user can plot multiple things at once in multiple views. Furthermore, each panel is *resizable* to any size the user would like!

Several areas still need improvement, such as fixing how the expansion panel ejects the plot off the renderable part of the card, the height responsiveness on the resize (only does it to width for some reason), and crashes on filter change with imshow (aggregation plot). Additionally, we need to except click events for the drag so that it doesn't interfere with zooming, selecting, or panning.

Added relay layout callback (Monday, Jan 8, 2024 - 13:47)

I added a relay layout callback on everything except skyplot and attempted min-maxing on zoom. During the commit squash, I added a relay layout callback specifically on skyplots.

Editor now returns error messages (Monday, Jan 8, 2024 - 13:47)

I modified the editor to return proper error messages, instead of just saying it failed. It uses the AssertionError's message and tells the user exactly why their expression is invalid.

Fixed skyplot min/max (Monday, Jan 8, 2024 - 13:45)

I fixed the minimum and maximum values for skyplots.

Fixed expeditor to catch 0 length expressions (Monday, Jan 8, 2024 - 10:28)

I fixed the expeditor to catch expressions which turn the dataset to 0 rows by doing a really long numpy check. This should probably be fixed in future,

Implemented download button (Saturday, Jan 6, 2024 - 11:44)

I implemented a download button, although it's currently unstable and may crash on higher subset sizes. It significantly impacts RAM usage (especially if you want a LARGE subset) and requires further optimization to store data locally on the server side while ensuring thread safety. Infuture, we need for the data to be stored as a "static file" to avoid such memory issues, but there could be issues with being threadsafe.

Added title, page title, and dark theme button (Saturday, Jan 6, 2024 - 07:52)

I added a title to the webpage, page title in the topbar, and dark theme button. However, the dark theme functionality does not work because it's not implemented in Solara at all.

Implemented skyplots (ra vs dec on projection) (Thursday, Jan 4, 2024 - 10:08)

I implemented skyplots, allowing for ra/dec and galactic longitude and latitude on aitoff and mollweide projections. Additionally, I added a relevant sidebar panel options, state variables, and plotting menu.

Attended Astra call on Jan 4 (Thursday, Jan 4, 2024 - 10:07)

I attended an Astra call on January 4, where I took feature requests and demo'd the app to the team. There were ideas about accessibility for font size and colorscale, and potentially around selection on scatter plots.

Fixed crash on expr filter set 2nd or later (Tuesday, Jan 2, 2024 - 09:33)

I fixed a bug that caused crashes when setting the filter to the dff object instead of the df object. Additionally, I coded my own version of the summary card partly in ipyvuetify reaction bits. I deleted relevant todo comments as well.

Implemented scatter hoverinfo to show ID (Friday, Dec 29, 2023 - 10:09)

I implemented scatter hoverinfo to show the SDSS_ID. It's easily changeable if we need to in future by the "hovertemplate" spec, which is really nice. Additionally, I adjusted todo formatting and added a component decorator to expreditor.

Implemented hist2d bin selection (Friday, Dec 29, 2023 - 09:39)

I implemented hist2d bin selection, allowing users to select a bin by clicking and setting the x/y limit of the plot to that bin. This can enable in-depth exploration if desired.

It's quite finicky, so I might remove it later, and just try to find out how to do bin selection.

Enabled bin edges retrieval on click (Sunday, Dec 24, 2023 - 11:13)

I made several changes, including adjusting the bin slider and adding a reset button to the UI. Additionally, the system successfully obtains the edges of the bin clicked on. There are todos remaining to make the plot edge changes and redraws correctly and to address sidebar redraws.

Moved log/flip card to bottom of sidebar (Saturday, Dec 23, 2023 - 21:28)

I fixed the position of the log/flip card to the bottom of the sidebar.

Implemented scatterplot hoverdata and started binselect (Saturday, Dec 23, 2023 - 21:14)

I implemented scatterplot hoverdata and started working on binselection, which is still a work in progress for the hist2d.

Added week 5 report (Saturday, Dec 23, 2023 - 21:14)

I made the week 5 report.

Implemented colorscale and UI updates (Thursday, Dec 14, 2023 - 11:14)

I implemented colorscale and UI updates, including dependencies-based render updates for plots. There might be a need for a thread-based render for the sidebar's pivot table.

Implemented scatter plot selections (Thursday, Dec 14, 2023 - 10:17)

I implemented scatter plot selections.

Fixed issue with catching categorical columns in scatterplot (Wednesday, Dec 13, 2023 - 11:27)

I fixed an issue with catching categorical columns in the scatterplot.

Implemented custom expression editor for the webapp (Wednesday, Dec 13, 2023 - 11:27)

I implemented a custom expression editor for the webapp, all written in Solara. This was the solution all along – don't try to bugfix the one Marteen Breddels wrote, write my own. I'm considering rewriting the pivot table as well, but I'm uncertain about it.

Added Solara repository for development work (Wednesday, Dec 13, 2023 - 09:40)

I added the Solara repository for development work into my thing, because of this very annoying expression editor, which will sometimes crash.

Added categorical data histogram, categorical warning, and other fixes (Monday, Dec 4, 2023 - 09:59)

I added a categorical data functionality to the histogram tab, and categorical data warnings for other views, and other fixes. Note that I modified Solara's source to fix the editor reset bug.

Fixed issue with hist2d y-axis flip (Monday, Dec 4, 2023 - 08:58)

I fixed an issue with the hist2d y-axis flip by changing the `px.imshow` origin option.

Attended meeting and worked on Nov 29 (Monday, Dec 4, 2023 - 08:54)

I attended a meeting and did some work, which was on a different machine and non-committed.

Added `vaex_jupyter` test (Wednesday, Nov 29, 2023 - 10:39)

I added a `vaex_jupyter` test. However, it doesn't work with Solara's `display/FigurePlotly` function, rendering it useless.

Implemented partial cross-filtering (Tuesday, Nov 28, 2023 - 09:43)

I implemented partial cross-filtering with Solara's undocumented filter cards and cross-filtering. I still need to implement selection cross-filtering on scatter plots (2D and 3D). Additionally, I fixed histogram log menu checkboxes, but they still don't work.

Added `test_dataframe_components.py` (Tuesday, Nov 28, 2023 - 08:58)

I added a new file, `test_dataframe_components.py`, for testing the undocumented components. I have mixed feelings about the plotting ones, but the summary and other components seem good.

Implemented aggregated binby statistic on histogram2d plots (Thursday, Nov 23, 2023 - 13:11)

I implemented aggregated binby statistic on histogram2d plots, which is now functional.

Added sidebar file, 3D plot, and bintype selection (Thursday, Nov 23, 2023 - 11:26)

I added a sidebar file, 3D plot, and bintype selection. Bintype selection is currently my biggest hurdle, as Vaex is behaving strangely with statistics.

Added plots, state, and dataframe submodules (Wednesday, Nov 22, 2023 - 10:02)

I added plots, state, and dataframe submodules, reorganizing everything for cleaner structure. I also changed `PlotState` to initialize not as `None` anymore to avoid potential issues.

Added `run.sh` script (Thursday, Nov 16, 2023 - 13:16)

I added a `run.sh` script, which runs the Solara application on a Starlette server.

Implemented functioning Vaex + Solara demonstration (Thursday, Nov 16, 2023 - 13:08)

I implemented a functioning Vaex + Solara demonstration. It can plot data in three different ways: scatter, hist, and hist2d (surface density). However, scatterplot performance is not optimal as it removes Vaex's advantages with out-of-core dataframes.

Fixed spec mask to use OR_MASK (Saturday, Nov 18, 2023 - 11:08)

I fixed the spec mask to use OR_MASK as suggested by Sean Morrison in the pull request thread. This change might be reverted in the future, depending on what the astronomers want.

Added bitmasks to Spectrum object outputs (Friday, Nov 17, 2023 - 11:47)

I added bitmasks to Spectrum object outputs. The Spectrum1D initializer now converts any 0 to valid values, assuming that zeroes in the bitmask indicate validity.

Updated .gitignore list (Friday, Nov 17, 2023 - 11:18)

I updated the .gitignore list, keeping .jupyter in case anyone else uses vim-jupyter during development.

Implemented individual identifiers and updated unit tests (Friday, Nov 17, 2023 - 11:17)

I implemented individual identifiers and updated some unit tests.

Deleted mpl_preamble.py (Friday, Nov 17, 2023 - 10:19)

I deleted mpl_preamble.py.

Fixed header fetch method to specutils standard (Friday, Nov 17, 2023 - 10:18)

I fixed the header fetch method to conform to the specutils standard. Now, it obtains the header from PrimaryHDU in the HDUList, and any data previously accessed through it has been removed as well.

Updated docustrings (Thursday, Nov 16, 2023 - 13:28)

I updated docustrings throughout the loader codebase.

Implemented unit tests with assertions (Thursday, Nov 16, 2023 - 09:37)

I implemented unit tests with assertions. However, I still need to write an MWM dummy file for the tests. Additionally, there's a new foobar variable check for the metadata to see if it gets loaded correctly.

Implemented unit tests on dummy data (Wednesday, Nov 15, 2023 - 14:26)

I implemented unit tests on dummy data, excluding MWM files. I need assistance with writing a BinTableHDU for MWM files, because for some reason, I can't figure it out.

Fixed header method to .get() and other minor fixes (Wednesday, Nov 15, 2023 - 14:25)

I fixed the header method to .get() and made other minor fixes.

Fixed handling of jdaviz specviz 1D in 2D array (Wednesday, Nov 15, 2023 - 11:20)

The jdaviz bug wasn't actually the NaN values. Andy argued that we should be able to keep them for science purposes. jdaviz DOES flag them out anyways, so I kept looking for something else. I fixed this issue in jdaviz because it didn't like a 1D flux in a 2D object – it gets confused and explodes. If the inner length of the 2D arrays is 1, I flatten the array. This fix lets us keep all zero and NaN flux points. Hooray!

Ignored demonstration and test files (Wednesday, Nov 15, 2023 - 09:58)

I added the demonstration and test files to the gitignore file.

Fixed Astropy units warning and warning format (Tuesday, Nov 14, 2023 - 13:24)

I fixed the Astropy units warning and changed the warning format to print.

All multiloaders functional (Tuesday, Nov 14, 2023 - 13:20)

I made all multiloaders functional. However, the apStar loader is not yet tested because the file is of length 1 with no visits. Additionally, the MWM loaders will skip any DATASUM=0 because Spectrum1D cannot be instantiated with zero data.

Fixed jdaviz nan and zero flux bug (Tuesday, Nov 14, 2023 - 12:59)

There's a very strange jdaviz bug, where it gets annoyed at nan and zero flux, and complains about dimensions in the flux values. The loaders mostly work, but jdaviz doesn't like *something* about the loaded data.

I fixed this bug by removing nan and zero flux points. I plan to open an issue on the jdaviz repository regarding this bug. The bug originates in the `x_min` `x_max` values used for the redshift slider for `line_lists` on nan and zero flux values in the Spectrum object.

Added test_implementation jupyter notebook (Tuesday, Nov 14, 2023 - 10:48)

I added the `test_implementation` jupyter notebook, although it is currently non-functional because of zero values in the x-axis. I also deleted `test_implementation.ipynb` for policy reasons – it would allow a user to see the currently generated data within the notebook.

Functioning loaders with refactoring and documentation (Tuesday, Nov 14, 2023 - 10:48)

I implemented functioning loaders with refactoring and documentation. This includes refactoring BOSS spec methods and MWM spec methods into single functions for simplicity.

However, the apStar multi-loader remains unimplemented for now, due to some issues. I need to check if files need to be cleaned of zero values and implement flux unit value fetches for spec and MWM files. Additionally, I need to check with the data team what versions of MWM files are needed for support – I'm presuming newest only.

They don't exactly work properly.

Partial implementation of loaders (Monday, Nov 13, 2023 - 12:34)

I partially implemented my loaders into the package.

The MWM loader is confirmed working. Remaining tasks include adding an HDU not-specified message, merging the MWM types into a single loader function, and confirming all other loaders work and adding them to `__all__`.

Added test_implementation.py (Monday, Nov 13, 2023 - 11:35)

I added `test_implementation.py` and plan to add the loaders directly into the package's functionality.

Identifier and documentation update (Monday, Nov 13, 2023 - 11:26)

Chore to update identifier and docs

SpecLite and other BOSS REDUX loader functionality (Monday, Nov 13, 2023 - 10:47)

I implemented SpecLite and other BOSS REDUX spectra loader functionality. Now, all BOSS spec can be loaded directly with the same underlying code, which required refactoring methods in the individual BOSS_spec loaders.

Added Solara Vaex prototype files (Sunday, Nov 12, 2023 - 15:52)

I added Solara Vaex prototype files: `scatter.py` and `re_export.py`. The `re_export` file is for creating a simple datafile for testing purposes (with $n = 1.1M$) from some old ApogeeNET data I had.

There are a few issues with the PlotlyResampler library, namely that its for time series data only. We would need to use a different aggregation method if we're set on using plotly (I am because its the only interactive backend that's easy to implement in solara).

Currently, I can only get Plotly to work for approximately 100K datapoints – anymore freezes the browser with upwards of 1.1GB of memory.

Plotly renders EVERYTHING in the front-end browser-side, which is why it leads to high memory consumption. I presume it has to do with the hover information as well – they’re distinct HTML objects, not just a raster layer.

I want to have an interactive, but also working app.

In theory, I could code an aggregator, but that seems quite out of scope. It might come up though.

Added mwmVisit and mwmStar loaders (Saturday, Nov 11, 2023 - 18:46)

I added mwmVisit and mwmStar datatype default loaders, and updated the demonstration notebook accordingly. I exported this output to a PDF to share with the collaboration. Additionally, I deleted some bad files which, if pushed, would violate the policy about sharing proprietary SDSS data – a graph would be available publicly.

Added BOSS specFull loaders (Saturday, Nov 11, 2023 - 09:47)

I implemented BOSS specFull loaders, allowing loading of BOSS specFull files. I also made other changes, such as commenting out some ipython autoreload commands which I was using to try and debug why the package wasn’t reloading – they never worked, the easiest way was to just restart the ipython shell.

On the specFull loaders, I have no idea what spAll, zAll, and zList are, but they’re non-spectra, so I don’t load them. I don’t think the InverseVariance property needs a unit, but I should stay peeled about this. The files also don’t have a flux unit within them, so they’re hard-coded as the SDSS standard.

Added astra_readers.py (Saturday, Nov 11, 2023 - 08:52)

I added astra_readers.py, pulled from the sdss/astra repository. This is Andy’s code which I’m using as a skeleton to implement the loaders properly, since these don’t work the way they want them to.

Added apStar/apVisit functionality and new helper funcs (Thursday, Nov 9, 2023 - 13:00)

I added apStar/apVisit functionality along with new helper functions: `__fetch__metadata` to perform a grab of common metadata, and `__fetch__flux__unit` to get flux unit from the given HDU and convert it to an Astropy unit object.

Added test.py for testing and fixed ordered method for apVisit (Thursday, Nov 9, 2023 - 10:30)

I added test.py for testing purposes and fixed the ordered method for apVisit.

Added sdss_v.py (Thursday, Nov 9, 2023 - 09:23)

I added sdss_v.py and started working on loaders.

Add the zora/valis docker to .gitignore (Thursday, Nov 10, 2023 - 13:10)

I added the zora-valis docker to the .gitignore file. Currently, I’ve been experimenting with it and trying to get it to work, but something in valis is causing a crash.

Initial project meeting and discussion of current loaders work (Thursday, Nov 9, 2023 - 12:29)

We had our first meeting to discuss the project and reviewed the current loaders work on the SDSS wiki. The project is simple and will allow me to get familiar with the SDSS datatypes, which is why I will do it before the parameter explorer. I also modified .gitignore to allow nested repositories for Astropy and Specutils.