



Rubin Observatory Data Preview 1

James Mullaney
The University of Sheffield



UK Research
and Innovation

What I will talk about today

- What Data Preview 1 is and isn't.
- What Data Preview 1 contains.
- How to access Data Preview 1.
 - Notebook tutorials on:
 - TAP access;
 - Butler access



In-prep DP1 release paper:
<https://rtn-095.lsst.io/>



What DP1 is...and isn't.

Data Preview 1:

- First on-sky data from the Rubin Observatory to be released for research;
- Was obtained using Rubin's **Commissioning Camera** (ComCam);
 - Same size, weight, detector type, filter complement etc. as the full camera, but containing only **9 detectors, instead of 189**.
 - Used to optically align the Simonyi Telescope and verify it can deliver the required data quality.
- Was obtained during the **48 nights** spanning 24th October - 11th December, 2024.

It is not:

- The data that was shown during the "First Light" event.
- *Exactly* what we expect for full Data Releases.

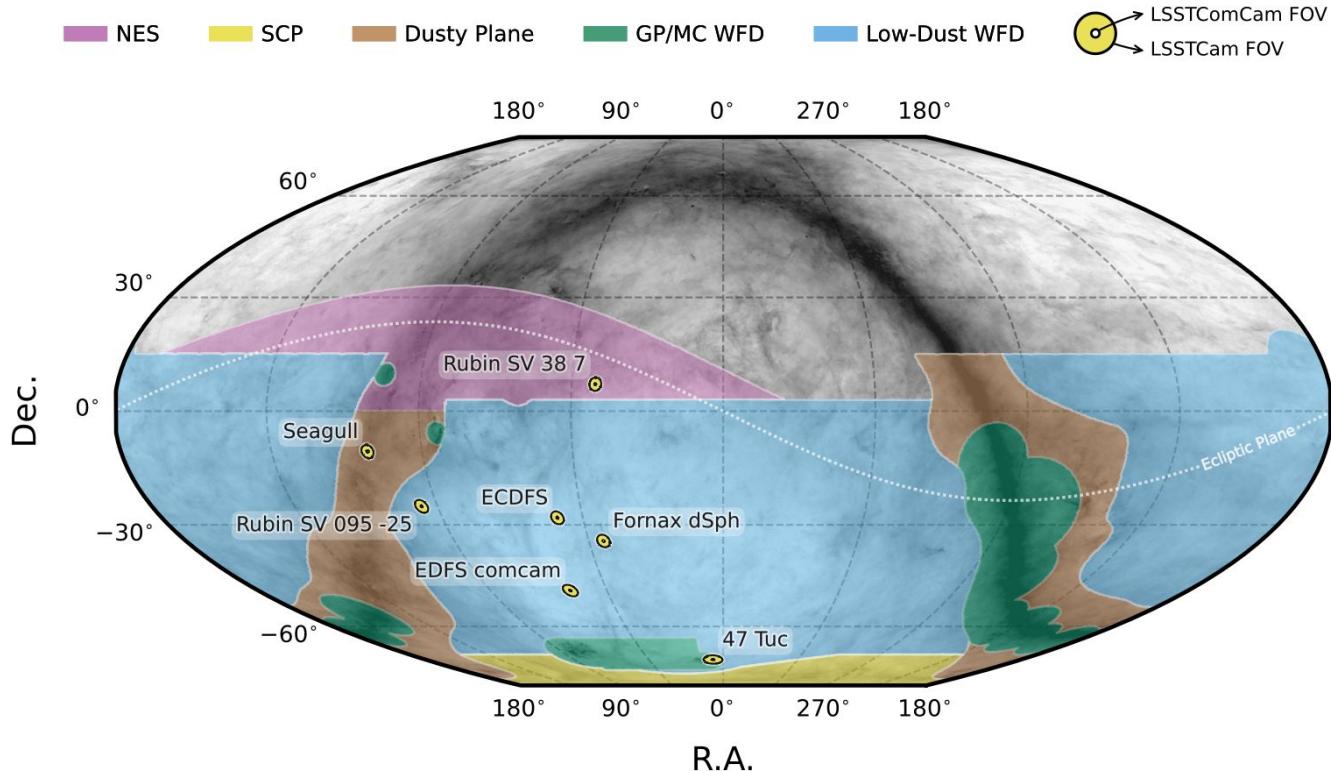
The ComCam Campaign

Targeted 7 fields:

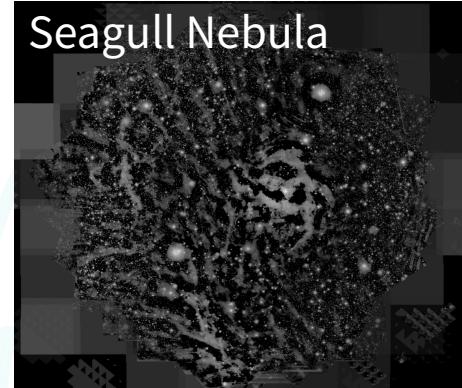
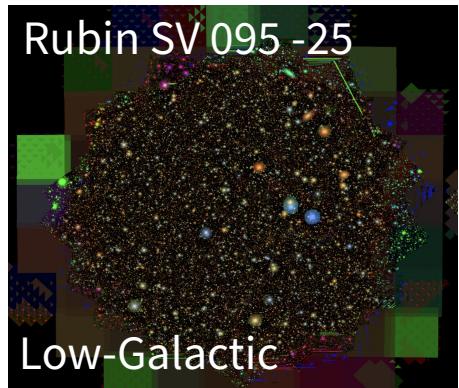
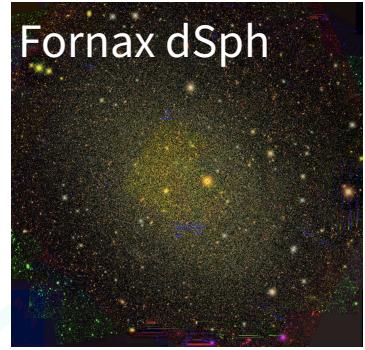
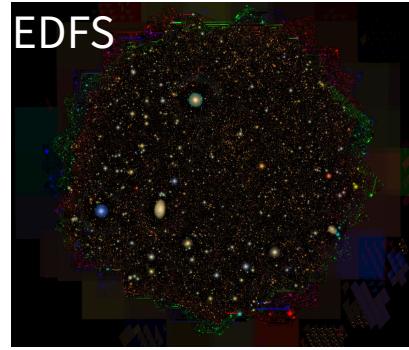
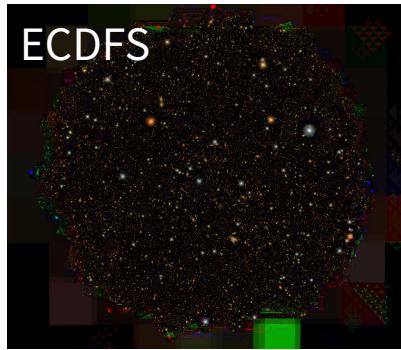
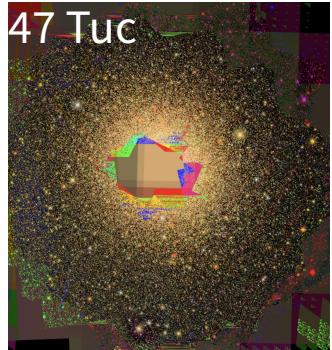
- 47 Tuc
- ECDFS
- EDFs comcam
- Fornax dSph
- Rubin SV 095-25
- Rubin SV 38 7
- Seagull

In all six bands:
u, g, r, i, z, y

Covered 15 sq. deg.



What the fields look like



What DP1 contains

Data in DP1 fall under five different broad types:

- Images
- Catalogues
- Maps
- Ancillary Data Products
- Metadata

Science Data Products

Aid investigation

Details of full contents at:
<https://dp1.lsst.io/>

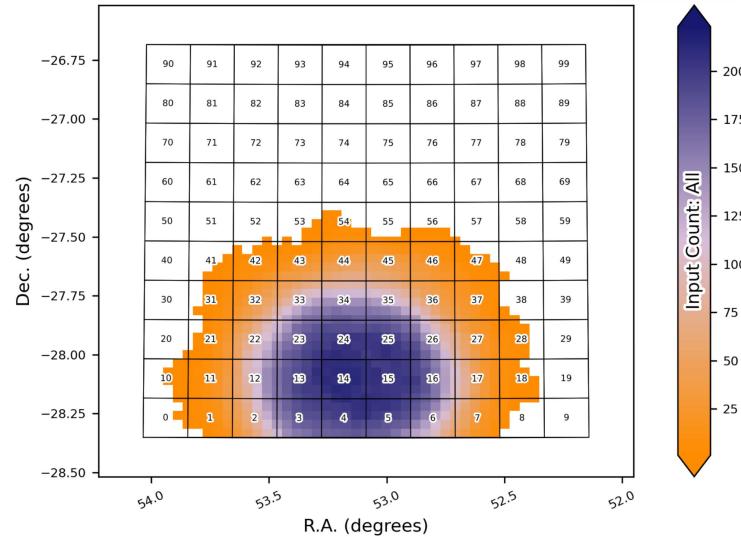


What DP1 contains - Images

- raw: Images directly from the camera, before any processing.
- visit_image: Processed & calibrated single epoch images (science, var, mask).
They are the lowest-level "scientifically useful" images.
- deep_coadd: Product of coadding visit_images taken with the same filter.
Selected to have PSF FWHM < 1.7 arcsec.
- template_coadd: Used to make difference_images.
The third of visit_images with the best seeing are selected.
- difference_image: From subtracting template_coadds from visit_images.
They show what has changed in flux.

Image Health Warnings!

- All processed images contain a mask plane.
 - Provides information about each pixel. Is it associated with a detection, a cosmic ray, a bad pixel....?
 - Make sure you understand and refer to the mask plane if you're working with image data.
- Particularly applies to coadds:
 - Many coadds only have partial coverage.
 - Regions without coverage don't necessarily contain - for example - zeros or NaNs.
 - But they **are** indicated with the NO_DATA pixel mask.
 - Ignore the mask planes at your peril!!!



What DP1 contains - Catalogues

- Source: Deblended 5-sig detections in `visit_images`. (46M sources)
- Object: Deblended 5-sig detections in `deep_coadds`. (2.3M objects)
- ForcedSource: Forced PSF photometry at object locations (269M entries)
- DiaSource: 5-sig detections in `difference_images`. (3.1M diaSources)
- DiaObject: Objects that diaSources are associated with. (1.1M diaObjects)
- ForcedSourceOnDiaObject: Forced PSF photometry at diaObject locations (197M)

Solar System:

- SSSource: DiaSources that are associated with known Solar System objects (5598)
- SSObject: Mapping between Rubin SS Object ID and the IAU designation.

Non-science/not-from-observations:

- CcdVisit: Information about the exposures. Pointing, PSF, etc. (16071 entries)
- Calibration: The catalogue used for astrom. and photom. calibration.

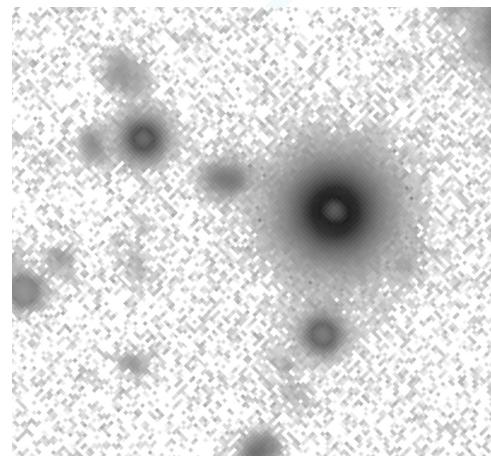
Catalogue Health Warnings!

Flags:

- Catalogues also contain flags.
- Indicate whether a detection/measurement is associated with a bad pixel, poorly-defined PSF etc.
- Make sure you understand and exclude appropriately flagged sources.
e.g., `~base_psfFlux_Flag`.

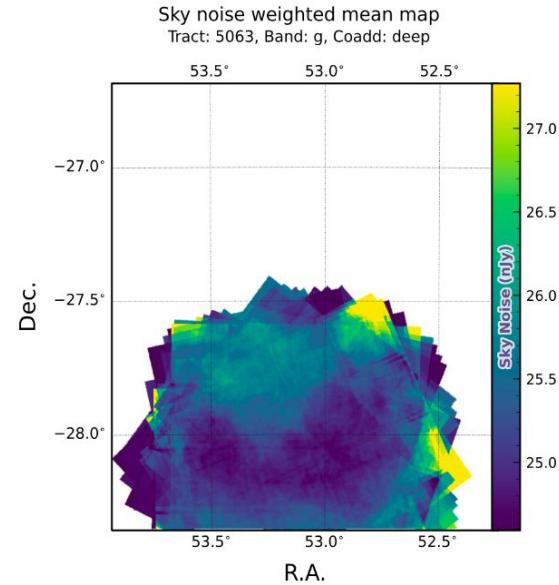
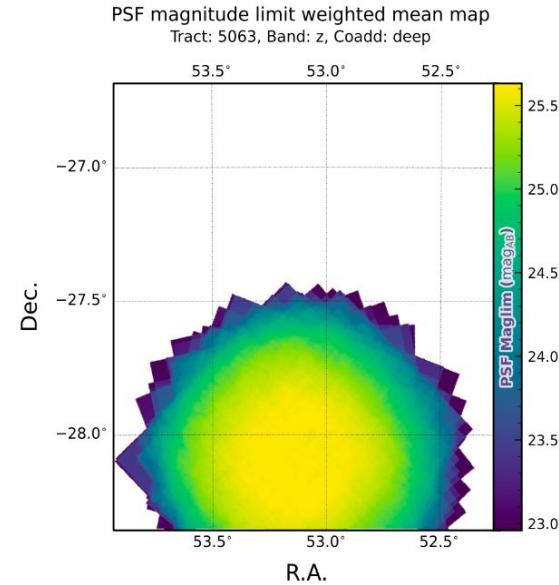
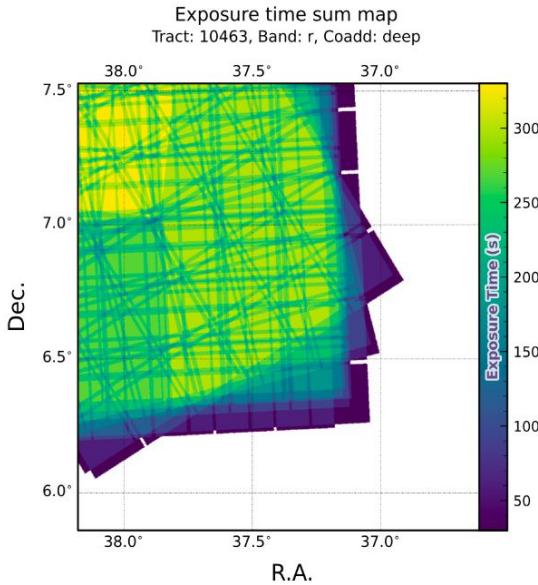
Lightcurves:

For lightcurves, we strongly recommend using the `psfDiffFlux` column in `ForcedSource`.



Maps

Maps are non-science-level images. Examples include:



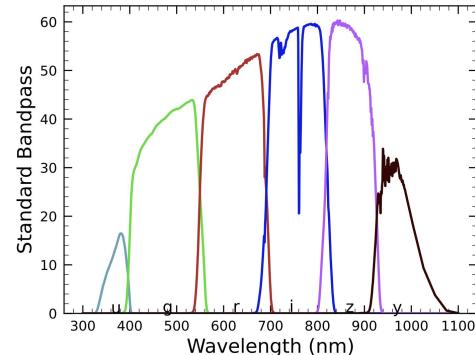
Metadata & Ancillary Data Products

Metadata: `visit` table contains data for each exposure.

Differs from `ccdVisit`: per exposure, as opposed to per exposure-detector.

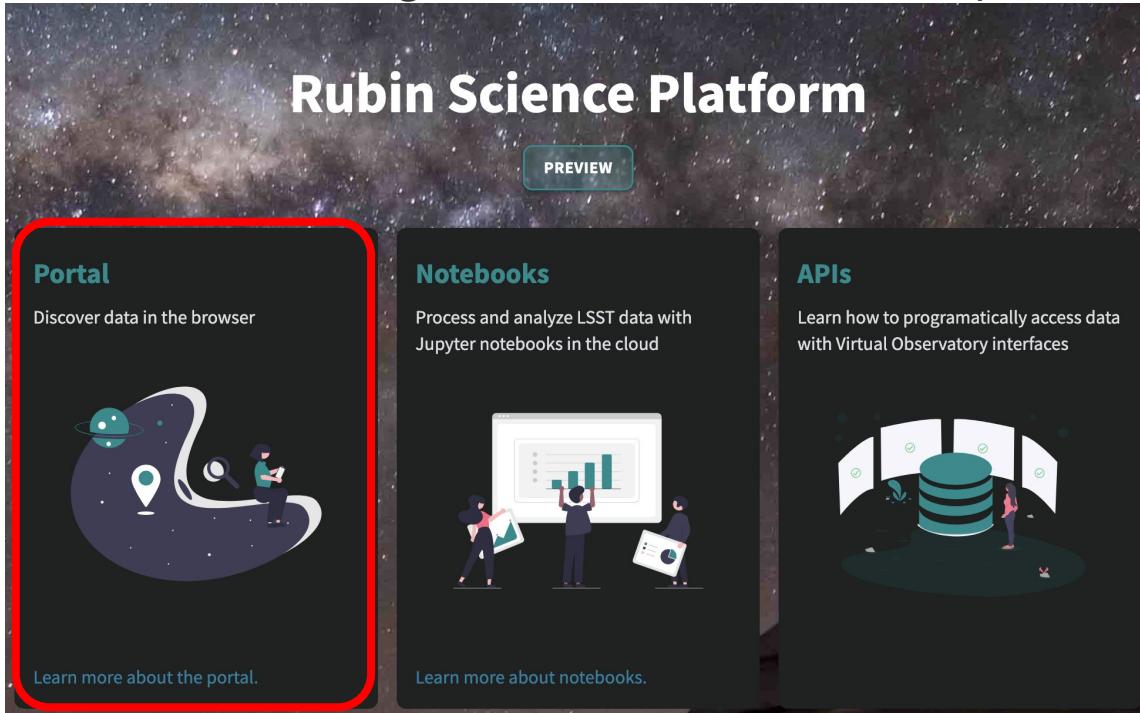
Ancillary Data Products:

- *Config files*: So users can know how the pipeline was configured.
- *log and metadata files*: Produced during processing, allows users to check why a certain processing failed, for example.
- *calibration data products*: Bias, flats, darks, brighter-fatter kernels, charge transfer inefficiencies...
- *whole system bandpass throughputs*...



Accessing the data

DP1 data is available **now** to Data Rights holders. Access is via <https://data.lsst.cloud>:





Video of accessing HiPS maps

Video of accessing Catalogues in Portal

Accessing the data

DP1 data is available **now** to Data Rights holders. Access is via <https://data.lsst.cloud>:

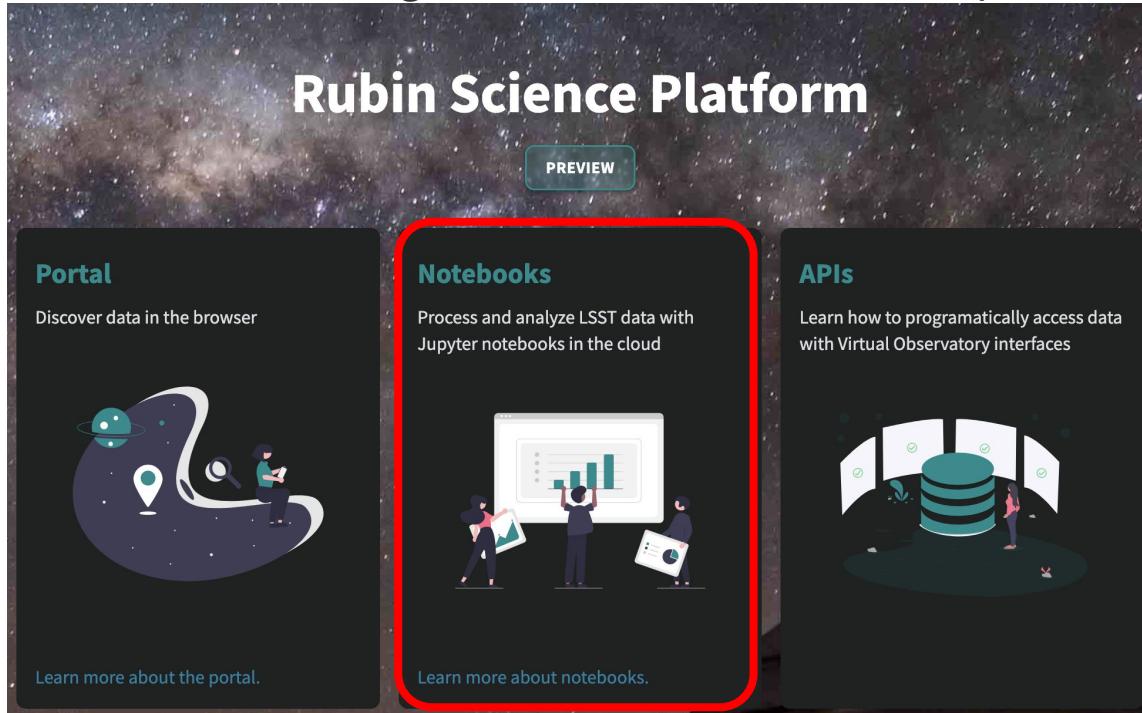


Table Access Platform (TAP)

- Query the same tables as the Portal Aspect, but in a notebook setting.
- It's the recommended way of programmatically accessing table data.
- ADQL - SQL with astronomy features - queries *all* objects, source etc. at once.
- Once searched-for and retrieved, data can be manipulated and analysed.
- It is faster - but less feature-rich and flexible - than using the Data Butler.

The Data Butler

- Use the Data Butler to retrieve pixel-level data for analysis.
- But it can retrieve table data, too.
- It's very powerful and feature-rich.
- Is more flexible than TAP, but less efficient for querying table data.
- Can be used to acquire broader types of information, such as tract and patch information.

Summary

- Data Preview 1 was made available to Data Rights holders on the 30th June 2025.
- DP1 is the first science-grade data to be released by the Rubin Observatory.
- It contains data from the ComCam campaign: 48 nights, 7 fields, 15 sq. deg.
- Excellent data quality in terms of photometry, astrometry, PSF, colours.
 - But we can't check *everything* - so we want to hear from the community.
- There are multiple ways to explore the data:
 - Portal Aspect to visualise images and catalogs;
 - Notebook Aspect to manipulate and analyse data:
 - TAP for straightforward catalogue searches;
 - Butler for image and more complex catalogue searches.
- <https://github.com/jrmullaney/LSSTAGN-DP1>

