

What is Data Science Studio?

Data Science Studio (DSS) from **Dataiku** is a complete Data Science software tool for data scientists, analysts and machine learning experts to perform data analysis and modelling more efficiently. DSS significantly shortens the time-consumed during data cleaning, model buildings and other statistical processes.

DSS enables direct and fast connection to the most common sources of data with strong integration capabilities. Analysts can leverage these smart data types to validate and transform the data in an automated way.

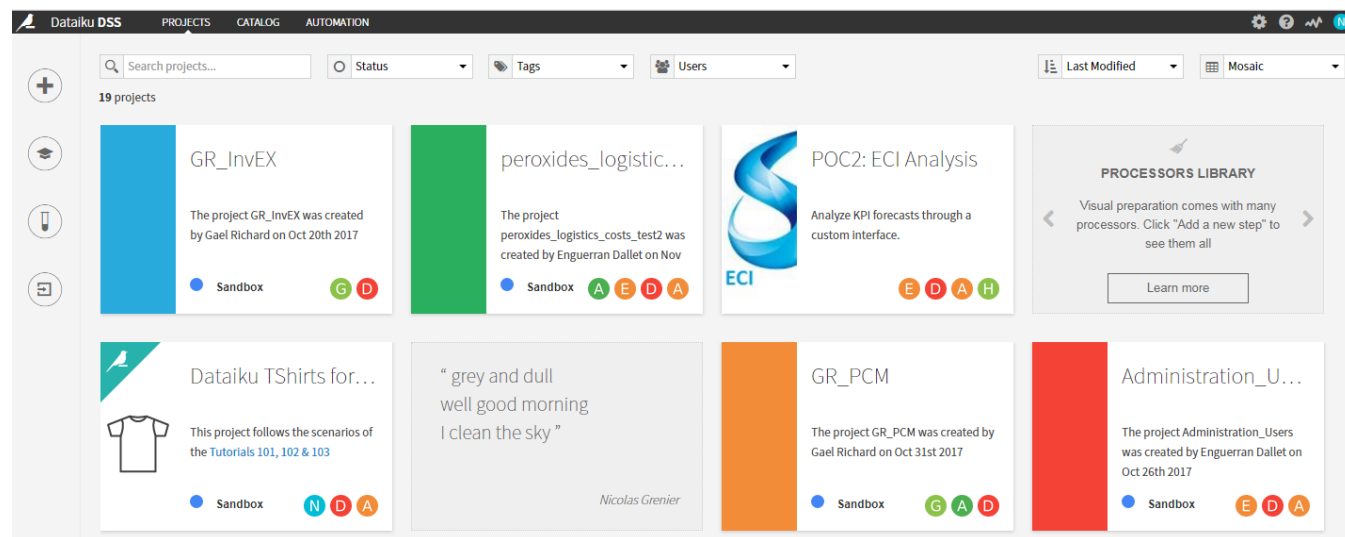
They can also perform more mundane tasks such as replacement, grouping, splitting, calculating, and others on DSS's interface which shows them an instant visual feedback of any operation.

Some Basic Concepts of Data Science Studio:

There are few main concepts that are important to understand the terms used to get more familiar with DSS.

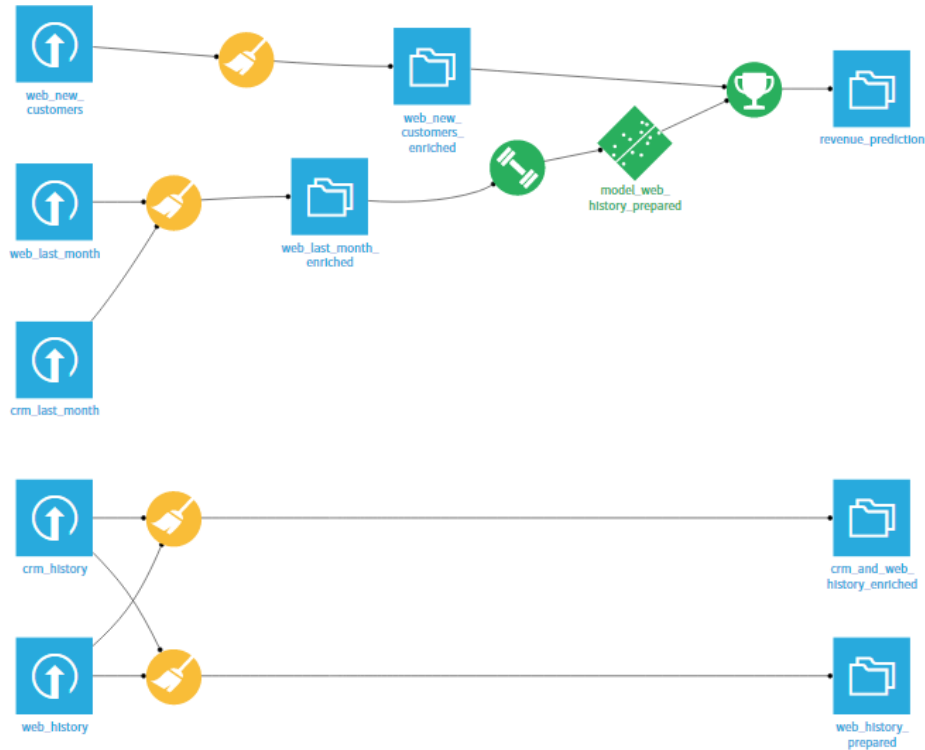
Projects

Each task in DSS is organized in individual projects to manage the data and associated tasks. The main dashboard is called as '*Universe*' and projects arranged will look like this:



Flow [blocked URL](#)

A DSS project is structured in the form of a flow. It visually represents a data pipeline and how datasets are



Datasets [blocked URL](#)

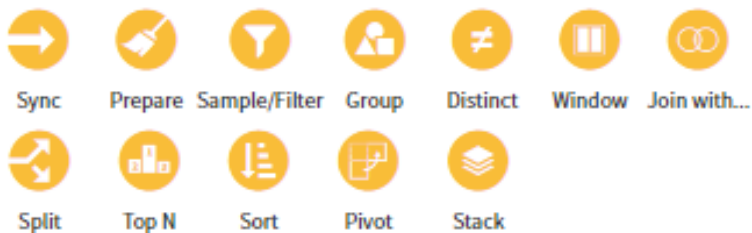
The dataset is the core object you will be manipulating in Data Science Studio. It is a series of rows with the same data structure.

Recipes

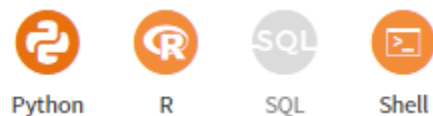
Any pre-processing or data manipulation on the datasets are managed using recipes. Recipes are the building blocks of your data applications. Each time you make a transformation, an aggregation, a join, ... with the Data Science Studio, you will be creating a recipe.

There are two types of recipes used widely in DSS:

Visual recipes: Provide basic manipulation functionalities like data cleaning, filtering, grouping etc.

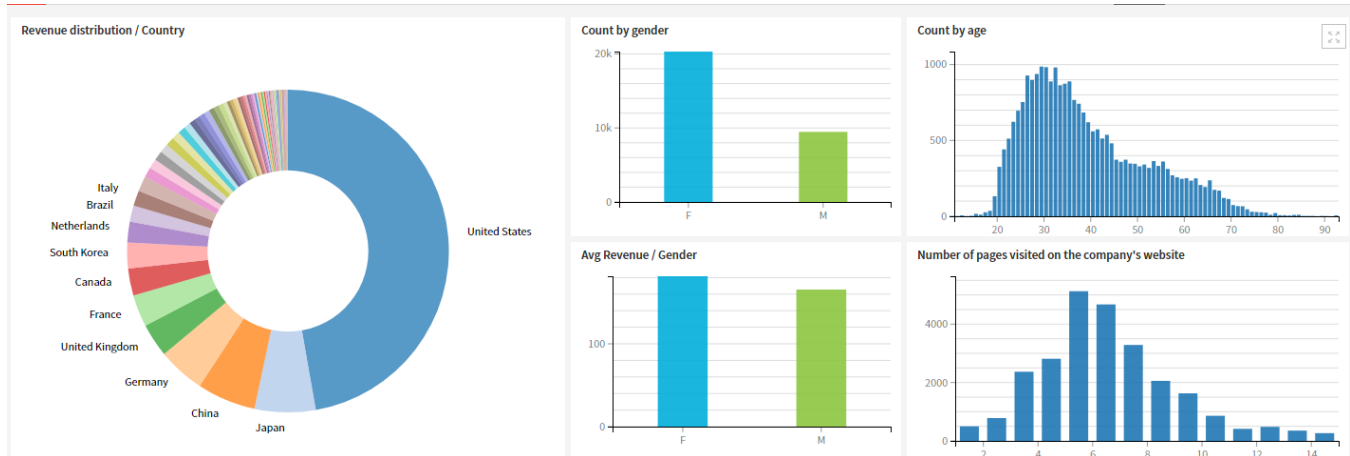


Code recipes: Used for integrating technical programming like R, Python etc.



Dashboard [blocked URL](#)

The dashboard communicates result and give insights based on the analysis performed on the datasets.



Analysis [blocked URL](#)

This provides visual analysis of the dataset prior to the implementation on the flow which helps to dive deep into the data directly.

Other concepts

Jobs: Every build on the dataset is recorded as jobs to keep track of activities in the flow

Scenarios: Helps in automating and scheduling the tasks in the flow

[blocked URL](#)

Lab - Notebooks: DSS allows to draft code in interactive programming environment to make the analysis easy and efficient

Web Apps: Users with Web coding skills can create advanced custom Web Apps using our dedicated editor and REST API

[blocked URL](#)

For more introduction on concepts of DSS, please navigate [here](#) .