

Sprint 1 :

Extracting the right kind of data that suits our problem:

The top screenshot shows the 'Select sources' dialog in IBM Cognos Analytics. The 'My content' search bar is empty. A list of sources is shown, including 'test_data.csv' (13/11/2022, 08:06). The 'Filter by' section on the right shows 'Type' filters (Folders, Packages, Files, Data sets, Modules) and 'Modified' filters (All, Today, Yesterday, Past week, Past month). The 'All' filter is selected under 'Modified'.

The bottom screenshot shows the 'Data module' view. The 'test_data.csv' file is selected in the left sidebar. The main table displays the following data:

Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available E...in Hc
1	318439	21	c	3	Z	3
2	318440	29	a	4	X	2
3	318441	26	b	2	Y	3
4	318442	6	a	6	X	3
5	318443	28	b	11	X	2
6	318444	23	a	6	X	3
7	318445	26	b	2	Y	2
8	318446	25	e	1	X	4
9	318447	23	a	6	X	4
10	318448	23	a	6	X	3
11	318449	10	e	1	X	4

Among the datasets we extracted 'test.csv' file seemed to be clean and detailed than others.