PDF is essentially a *container* that describes a document's layout and content in a way that can be understood by any device that can read the file. This ensures that the document is represented in the same way on any kind of device. This PDF *container* contains objects like pages, fonts, images and other metadata, that all get cross-referenced/linked together within a table. By reading the table and creating relationships between objects we can recreate the document.Here's an overview of content types we might find inside a PDF file

Category	Examples
Text content	Streams with font references, character codes
Vector graphics	Shapes, lines, bezier paths
Images	Raster data (JPEG, JPEG2000, CCITT, etc.)
Fonts	Embedded or referenced
Metadata	Info dict, XMP XML
Interactive content	Forms, annotations, links
Structure / Tags	Logical document structure for accessibility
Embedded attachments	PDFs, images, or arbitrary files
Scripts / Actions	JavaScript actions (rare, used in forms)

Let's also add an image here: