

Amazon Textract Overview

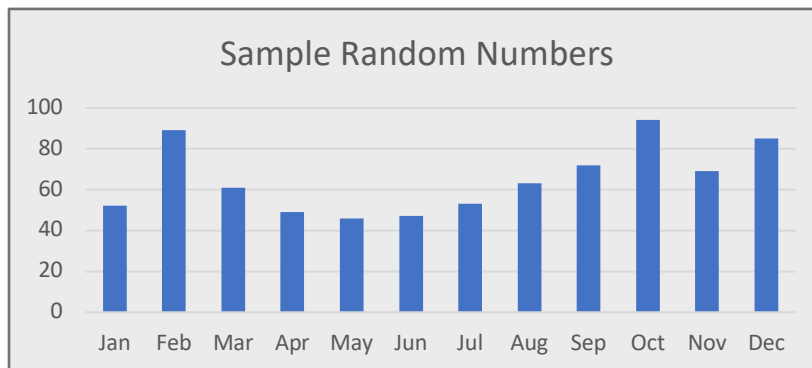
Amazon Textract is a fully managed machine learning service that automatically extracts printed text, handwriting, and other data from scanned documents that goes beyond simple optical character recognition (OCR) to identify, understand, and extract data from forms and tables.



Many companies today extract data from scanned documents, such as PDF's, tables and forms, through manual data entry (that is slow, expensive and prone to errors), or through simple OCR software that requires manual configuration which needs to be updated each time the form changes to be usable.

To overcome these manual processes, Textract uses machine learning to instantly read and process any type of document, accurately extracting printed text, handwriting, forms, tables and, other data without the need for any manual effort or custom code.

With Textract you can quickly automate manual document activities, enabling you to process millions of document pages in hours. Once the information is captured, you can take action on



it within your business applications to initiate next steps for a loan application, tax document, enrollment form or medical claims processing. Additionally, you can create smart search indexes, or add in human reviews with Amazon Augmented AI to review nuanced or sensitive data.