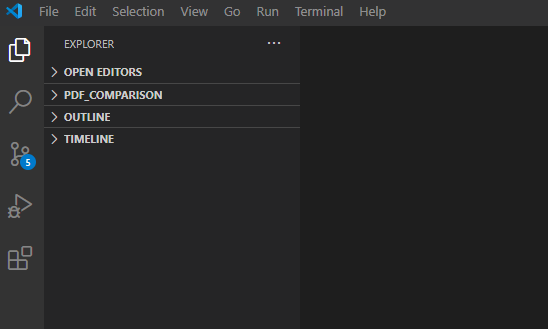
# **PDF Sample Document**

|  |  |
| --- | --- |
| **Sample Date:** | October 2020 |
| **Prepared By:** | Sample Client |
| **Created and Tested Using:** | * Framework 1.1 * Additional Framework 2.1.0 |

## **Overview**

This sample consists of a simple PDF document containing images, tables and bodies of text for use in future comparisons.

Each page features several different elements, which will be re-arranged and taken out of various versions to make appropriate comparisons. Below is an image:



## **The Comparison:**

* The comparison process comprises several parts
  1. Identify the file type of each document and extract the contents to a string using the appropriate method
  2. Pre-process the string for each document to remove any line-breaks and unusual characters
  3. Compare the strings for a perfect match
  4. If no match, separate each string into a list of sentences to identify the differences