Objectives

- **to design and automate a process that works with Excel and PDF files.**

Task	Description
[individual/team] **A. Excel and PDF Activities** **[due: Lab06]**	• Design and automate a process that o <u>extracts</u> data from (at your choice): ▪ .txt files, .csv files, etc; ▪ **PDF** file(s) (see **Lecture 10**) ▪ web pages using **data scraping** or **screen scraping**; o <u>stores</u> it into an **Excel** file(s) (see **Lecture 09**) and <u>performs processing</u> on it. • Use *various activities for each type of file* you work with. • E.g., consider various invoices given as PDF files. Extract relevant data (date, amount, client name, etc.) and build an Excel file with the extracted data. Filter and/or sort the data based on a given criteria; save the result into a different Excel file.

Recommended readings:

- from the **RPA Developer Foundation** learning plan skim the followings lessons:
  - **DataTables and Excel Automation with Studio (v2021.10)** (3h15min)
  - **PDF Automation in Studio (v2021.10)** (1h30min).