

# TICKETS DATA AGGREGATOR

## Ultimate C# Masterclass Assignment

### Overview

This application reads the content of PDF files from a given folder. Those files contain cinema tickets, and the application should aggregate them in a single \*.txt file.

Console App

### Links:

<https://github.com/UglyToad/PdfPig> - the PdfPig library that can be used to read text from PDF documents

### Main application workflow

The application is given a path of a directory with the ticket PDFs. The sample PDFs can be found in the resources of the video containing the description of this project.

The application should read the content of each PDF ticket, extracting the data about the title, date, and time from each ticket. Each PDF can contain multiple tickets. Each file is always one page long. See **Sample Ticket PDF** for reference.

The dates and times of the show can be formatted differently for various cultures. A culture can be recognized by the domain of the web page address in the ticket. See **Sample Ticket PDF** for reference.

The application shall support the following cultures:

- en-US (.com domain)
- fr-FR (.fr domain)
- ja-JP (.jp domain).

The application should save the data of all tickets in a single text document. The dates and times in this document should be formatted according to invariant culture. See **Sample Result File** for reference



Finally, the app prints:

Result saved to **PATH\_TO\_FOLDER\_WITH\_TICKETS\aggregatedTickets.txt**  
Press any key to close.

# Sample Ticket PDF

This ticket uses the en-US culture to format the date and time. Please notice the address - [www.ourCinema.com](http://www.ourCinema.com), which can be used to recognize that this ticker was sold in the US.

Thank you for choosing our cinema! Here are your tickets:

|   |   |
|---|---|
| Title: Blade Runner 2049<br>Date: 1/2/2023<br>Time: 7:30 PM |  |
| Title: Blade Runner 2049<br>Date: 1/2/2023<br>Time: 7:30 PM |  |

Visit us: [www.ourCinema.com](http://www.ourCinema.com)

On the other hand, this ticket was sold in Japan. It uses a different date and time format. The domain is **.jp**.


Thank you for choosing our cinema! Here are your tickets:

|  |   |
|--|---|
| Title: Knives Out<br>Date: 2024/05/08<br>Time: 17:45 |  |
|--|---|

Visit us: [www.ourCinema.jp](http://www.ourCinema.jp)

## Sample Result File

The data from all tickets should be aggregated in a single \*.txt file. Each ticket should occupy a single line. The date and time should be formatted according to invariant culture.

 aggregatedTickets.txt - Notepad

| File              | Edit | Format | View       | Help  |
|-------------------|------|--------|------------|-------|
| Blade Runner 2049 |      |        | 01/02/2023 | 19:30 |
| Blade Runner 2049 |      |        | 01/02/2023 | 19:30 |
| Knives Out        |      |        | 05/08/2024 | 17:45 |
| The Lion King     |      |        | 07/02/2023 | 14:15 |
| The Lion King     |      |        | 07/02/2023 | 14:15 |
| The Lion King     |      |        | 07/02/2023 | 14:15 |
| The Lion King     |      |        | 07/02/2023 | 14:15 |

Each time the application is run, it should overwrite the content of this file. This file can be saved in the same folder in which the tickets are stored.

## PDF-reading library

Any library can be used to read the text from PDF documents (assuming its license allows it).

One of the libraries that can be used is **PdfPig**:

<https://github.com/UglyToad/PdfPig><https://github.com/UglyToad/PdfPig>

It can be installed via NuGet. It has clear and simple documentation and is easy to use.