

**Investintech.com Inc.
PDF-to-Excel
Command Line Tool
User's Guide**

June 14, 2007
<http://www.investintech.com>

Copyright 2007 Investintech.com, Inc. All rights reserved

Adobe is registered by Adobe Systems Incorporated
Acrobat is registered by Adobe Systems Incorporated
Postscript is registered by Adobe Systems Incorporated
Access™ is registered by Microsoft Corporation
Excel is registered by Microsoft Corporation
Visual Basic is registered by Microsoft Corporation
Visual C++ is registered by Microsoft Corporation
Visual C# is registered by Microsoft Corporation
Visual J# is registered by Microsoft Corporation
Visual Studio is registered by Microsoft Corporation
Win32 is registered by Microsoft Corporation
Windows is registered by Microsoft Corporation
Windows NT is registered by Microsoft Corporation
Windows Vista™ is registered by Microsoft Corporation

Copyright 2007 Investintech.com Inc. All rights reserved.

Contents

Preface	1
About This Documentation	1
Typographical Conventions Used in This Document	1
Getting More Information	1
Customer Service and Technical Support	2
Fax and Mailing Address.....	2
PDF-to-Excel Command Line.....	3
What is Command Line?.....	4
What is PDF-to-Excel Command Line?.....	4
Installation Instructions for PDF-to-Excel	4
System Requirements	4
Starting the Installation.....	5
Using PDF-to-Excel	11
Using PDF-to-Excel from Command Prompt	11
Conversion type	12
Schedule	13
Template.....	13
Range	14
Input file(s).....	14
Output file(s).....	14
Examples of Using PDF-to-Excel	14
Example #1	14
Example #2	15
Example #3	15
Example #4	16
Troubleshooting	16
Index	18

Preface

- Welcome to Investintech.com Inc. PDF-to-Excel Command Line. This tool provides technologies for converting PDF files into file formats recognized by Microsoft Excel. This user's guide provides the basic information needed to use the command line tool.

About This Documentation

This documentation describes the purpose, installation and usage of the PDF-to-Excel command-line program that allows users to convert PDF files into various formats (Excel, RTF).

Typographical Conventions Used in This Document

The following table describes typographical and naming conventions used in this document:

<i>Font</i>	<i>Used for</i>	<i>Examples</i>
monospaced	Expected input from the user	PDF-to-Excel -help
	Paths and filenames	c:\investintech\PDF-to-Excel.exe
	Source code	MessageBox("ABC");
monospaced blue	Source code keywords	if
monospaced green	Source code comments	//this is a comment
monospaced bold	Method name	The AboutBox() method
'Single quoted bold'	Elements of the user interface like buttons, icons etc.	click on 'Start' button, press 'Enter'
blue underlined	Internet links	http://www.investintech.com

Getting More Information

Investintech.com Inc. conducts research in data conversion technology and develops data conversion products utilizing results of the research. More information about our company is available at the company's website <http://www.investintech.com>.

Customer Service and Technical Support

Investintech.com Inc. strives to provide the best possible technical support to its customers and prospective customers. If you would like personal assistance, please feel free to call us or send e-mail to Customer Service or Technical Support. We are available by phone during regular business hours and in the vast majority of cases, we will return our e-mail with an answer the same day it is received.

	<i>Customer Service</i>	<i>Technical Support</i>
Telephone:	+1 416 920 5884	+1 416 920 2539
E-mail:	cs@investintech.com	techsupport@investintech.com
Hours:	Our Business Hours are Monday to Friday 9am-6pm (Eastern Time GMT-5:00).	

Fax and Mailing Address

Mailing Address: Investintech.com Inc.
410-96 Spadina Avenue
Toronto, ON
M5V 2J6 Canada

Fax: +1 416 920 5848

Investintech.com Inc. Software Development Kit User's Guide

PDF-to-Excel Command Line

What is Command Line?

Microsoft Windows operating system is graphical user interface (GUI) operating system – commands are usually given by using mouse to interact with graphical elements on screen and results are visible through changes on graphics drawn on computer screen. Microsoft DOS is command line interface operating system – commands are typed on keyboard, executed when ‘**Enter**’ key on keyboard is pressed and results are visible in text mode on screen.

While GUI is generally better way of using computers, there are few cases where command line is much better solution for particular need – like when we want to repeat same command over and over again.

In order to get command line interface window in any version of Windows, click on ‘**Start**’ (usually at lower left side of screen), point to ‘**All Programs**’, then point to ‘**Accessories**’ and click on ‘**Command Prompt**’ (or ‘**MS DOS Prompt**’ in older versions of Windows).

What is PDF-to-Excel Command Line?

PDF-to-Excel Command Line is command line, shell or batch version providing a subset of features found in Able2Extract. It may be used to automate the conversion of multiple documents.

The demo version of PDF-to-Excel is available, free of charge, for download from our company website, <http://www.investintech.com>

Installation Instructions for PDF-to-Excel

Below you will find step-by-step instructions to install PDF-to-Excel Command Line on your system. You may want to print this instruction sheet for reference before beginning to install PDF-to-Excel.

System Requirements

The minimum computer system resources required to install and run PDF-to-Excel are:

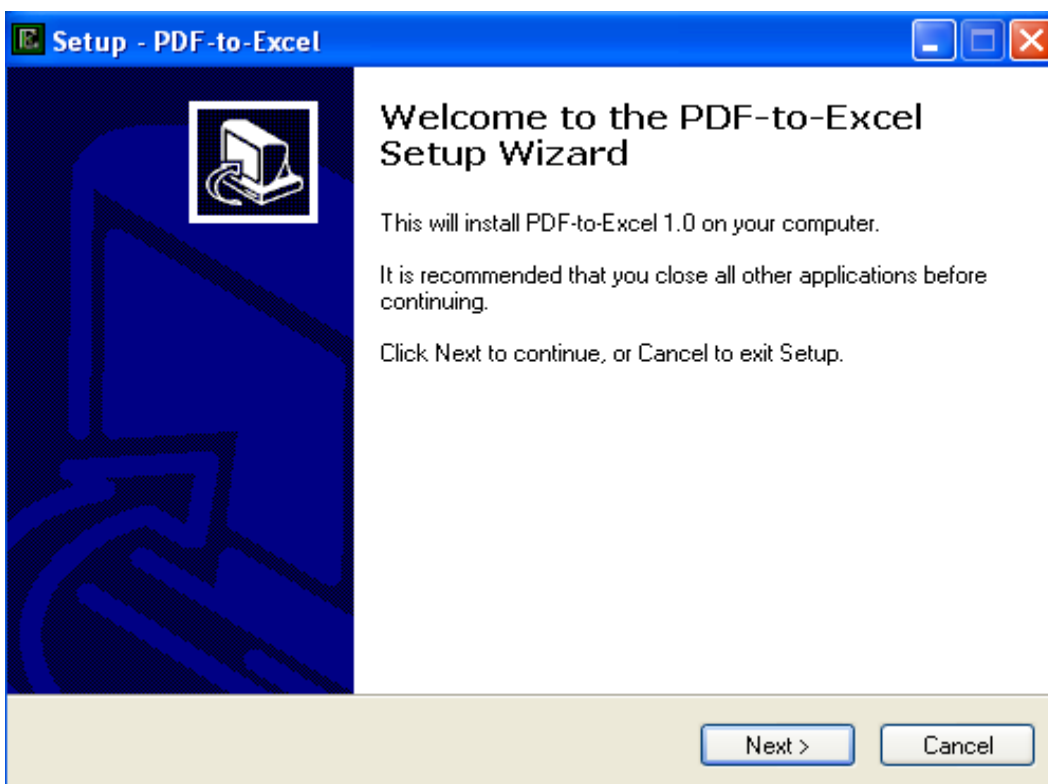
- Microsoft Windows 98/Me/2000/XP/2003
- 64 MB of RAM
- 4 MB of available hard disk space

In order to use PDF-to-Excel comfortably, it is recommended that your computer system meet the following requirements:

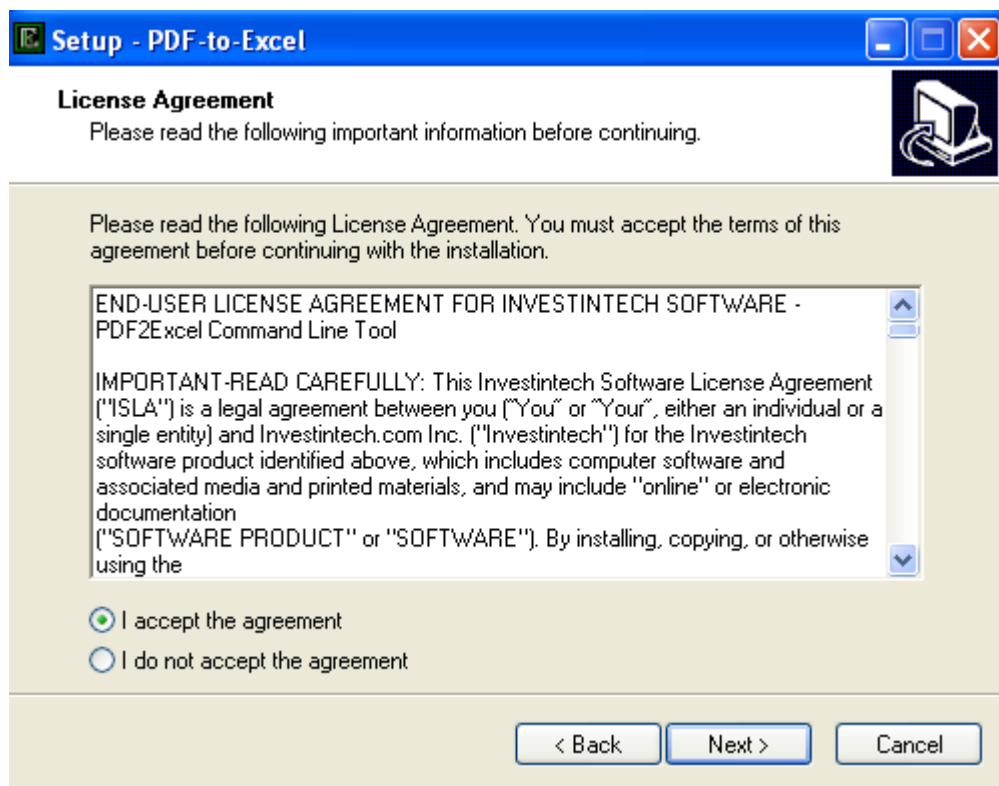
- Microsoft Windows 2000/XP/2003/Vista
- 512 MB of RAM
- 100 MB of available hard disk space

Starting the Installation

Start the installation by launching the `InstallPDF2Excel.exe` file, e.g. by double clicking it in the Windows Explorer. After that you should see a window that looks like this:

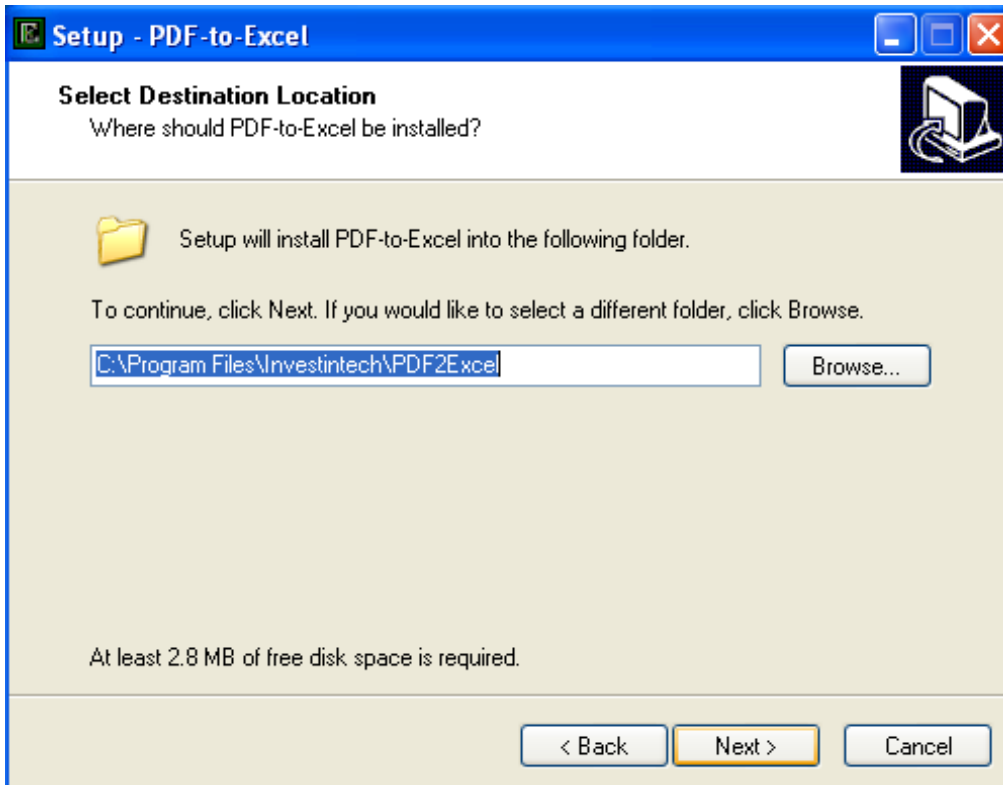


Click the **'Next'** button to continue installation. You should see the next window:

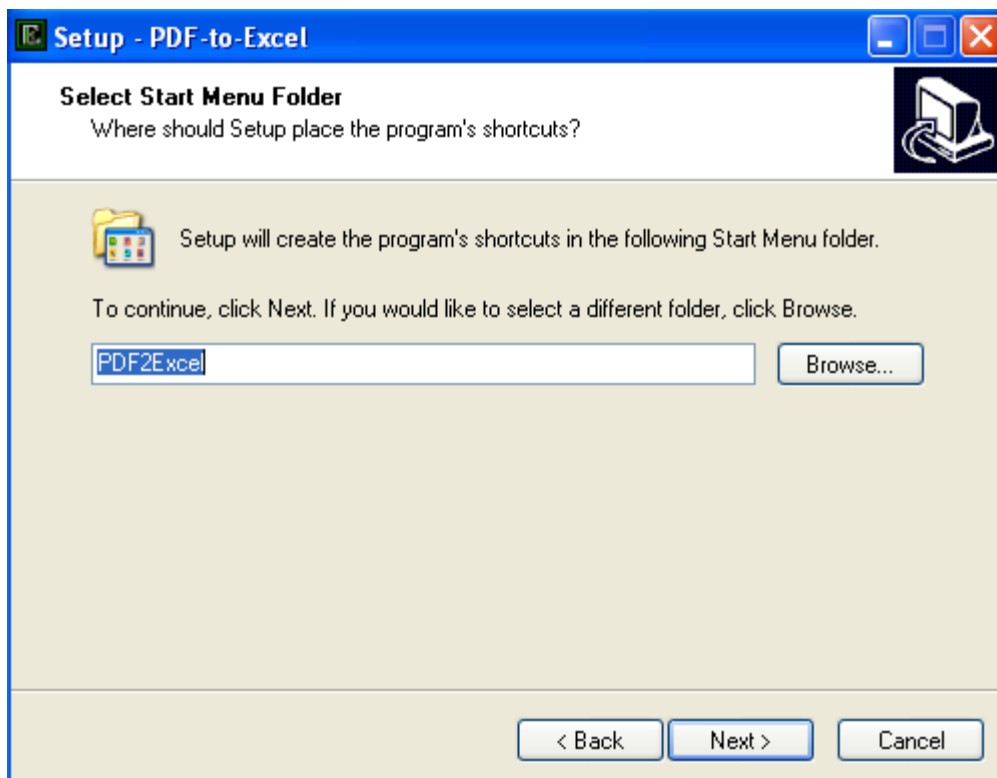


Please read the Investintech Software License Agreement carefully and thoroughly. If you wish to continue with installation, choose **'I agree with the above terms and conditions'** and click **'Next'**. Otherwise, click **'Exit'** and installation will terminate. If you clicked **'Next'** you should see the next window:

PDF2Excel Command Line

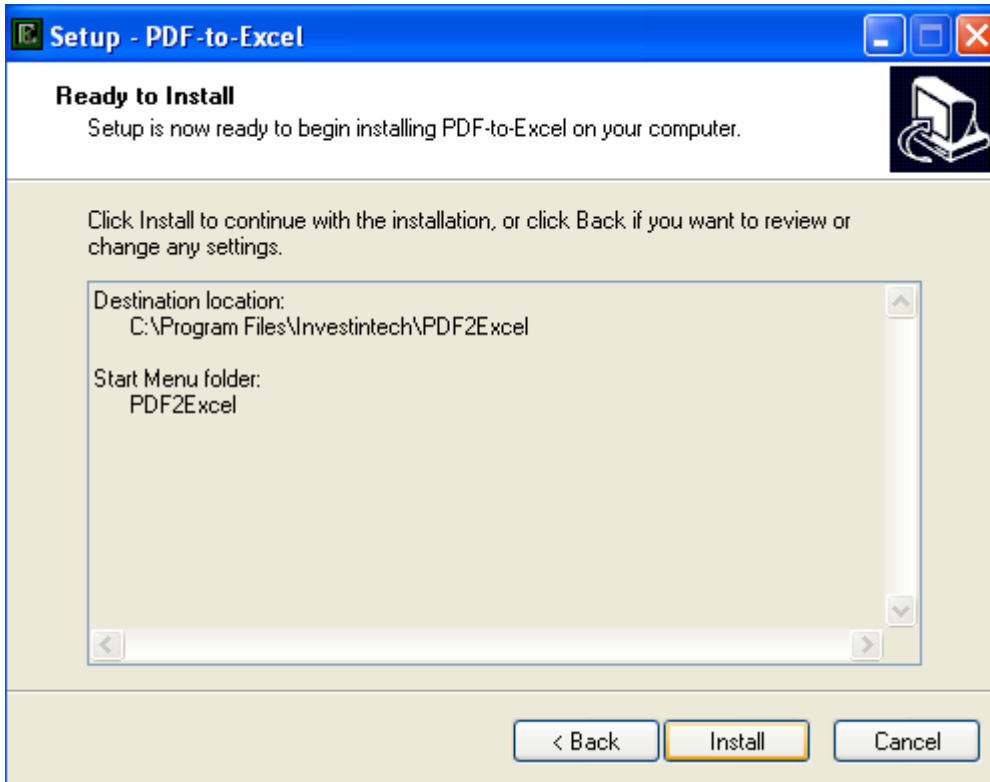


On this screen you can choose the installation directory for the PDF-to-Excel software. The default installation folder will be contained in the `Program Files` folder on your system disk. It is possible to change the location of installation folder either by typing its path directly or by clicking the 'Browse...' button. If the destination folder (directory) you have chosen does not exist, the installation program will create it. Click the '**Next**' button to continue the installation.

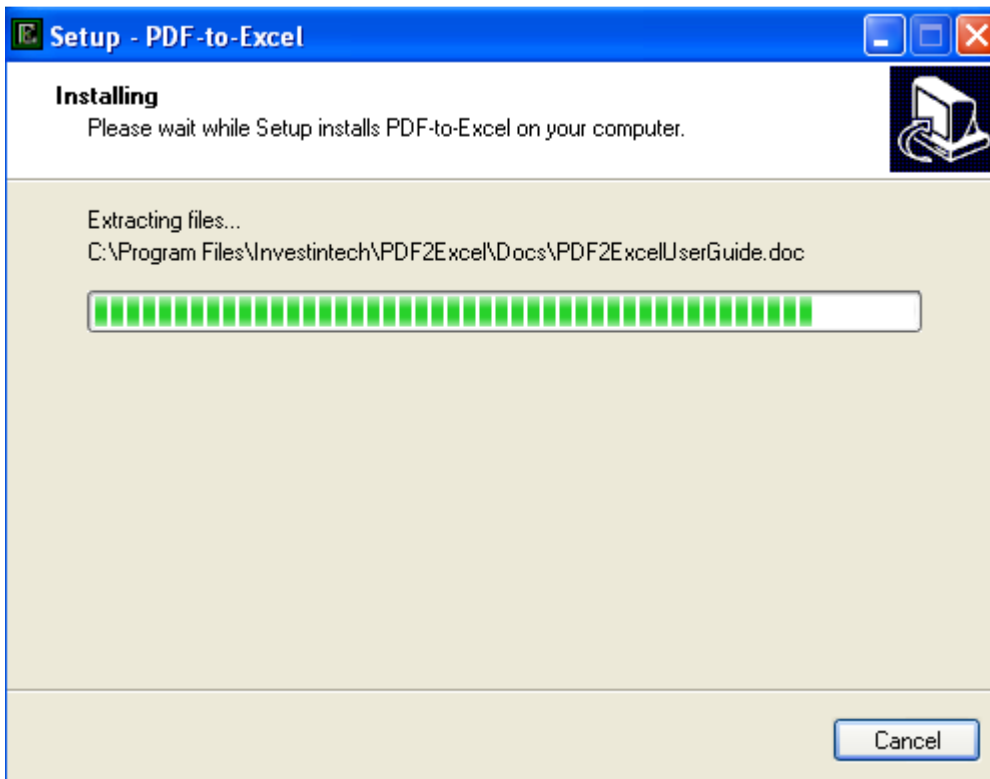


This screen prompts you for the for the start menu folder containing shortcuts to PDF-to-Excel. You may accept the default folder name, type a new folder name, or browse existing folders by clicking the “Browse...” button. Click the ‘Next’ button to continue the installation.

PDF2Excel Command Line

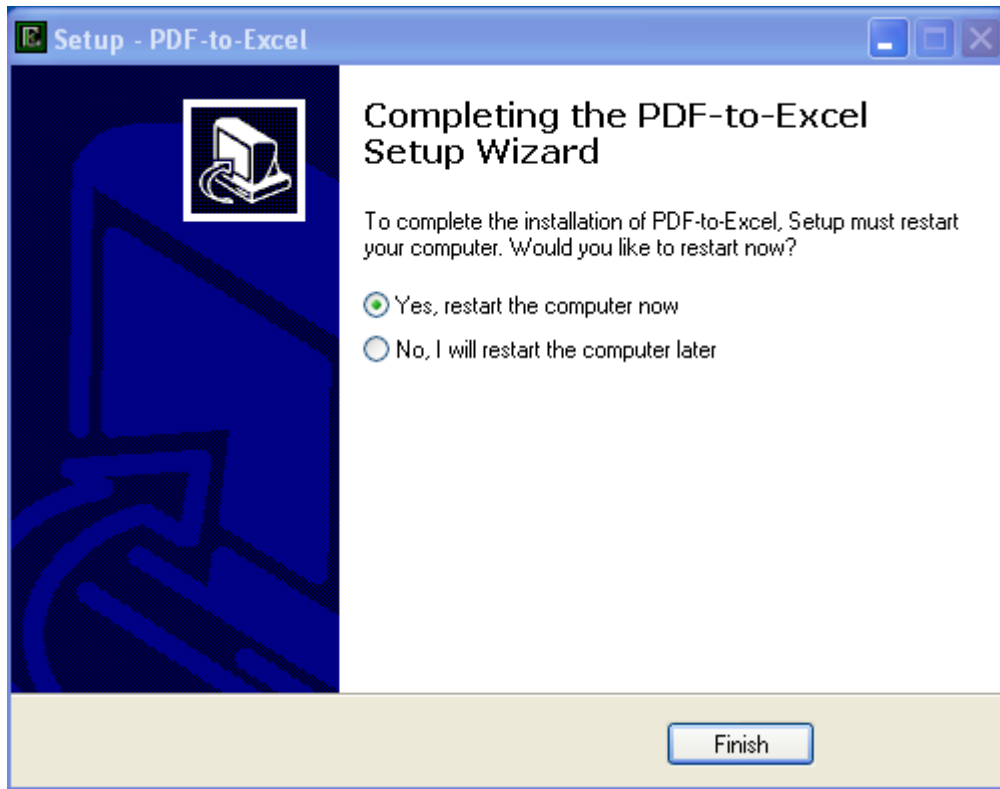


Clicking the **'Install'** button will start the installation.



PDF2Excel Command Line

After a very short period of time, usually about 30 seconds, you should see the next window:



The installation is now finished. Click '**Finish**' to finish the installation.

Using PDF-to-Excel

This section provides detailed information about using PDF-to-Excel software. The following topics are covered:

- using PDF-to-Excel from command prompt
- using PDF-to-Excel for batch conversion

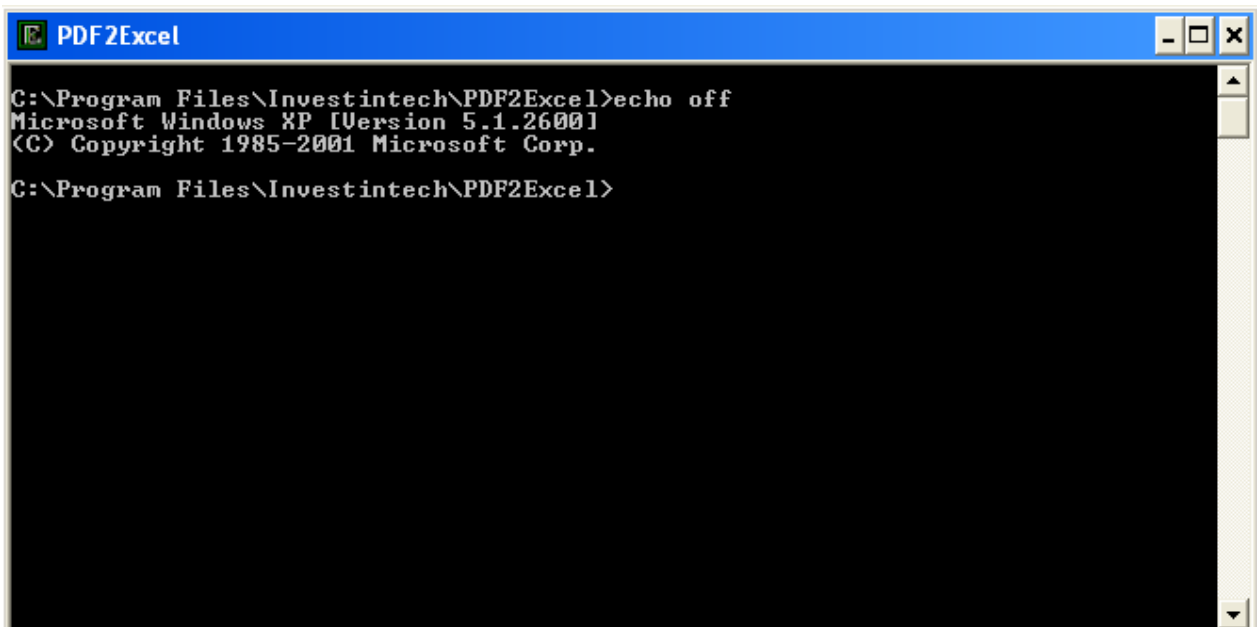
Using PDF-to-Excel from Command Prompt

Prior to using PDF-to-Excel, you must open a command prompt window. To open the command prompt window:



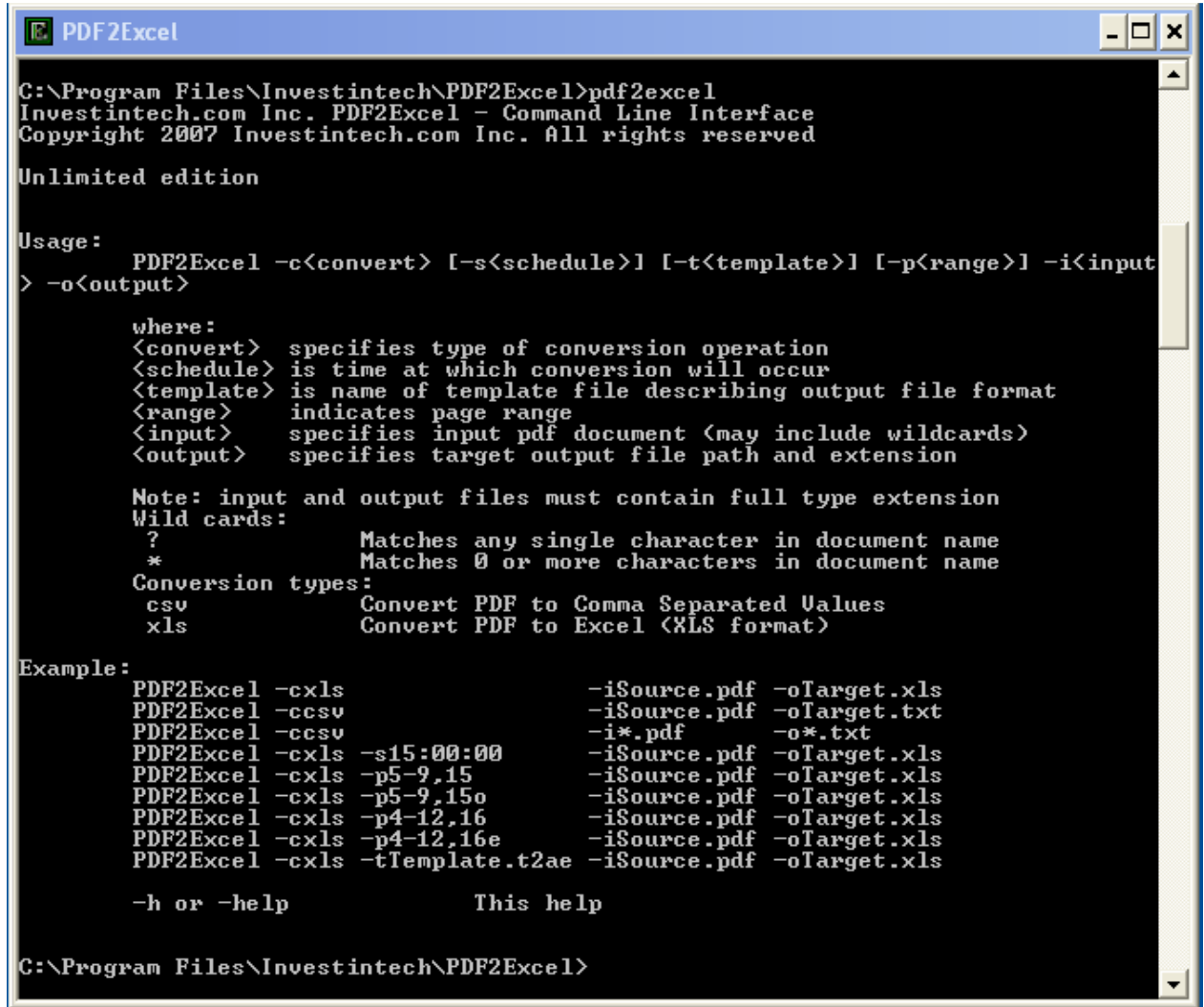
- Click the 'Start' button, then point to 'All Programs', then point to 'PD-to-Excel' and finally point and click on the 'PDF2Word' menu item.

The window titled "PDF2Excel" should open, as shown.



PDF2Excel Command Line

To display the PDF2Word help message, simply type PDF2Word and press 'Enter'.



```
C:\Program Files\Investintech\PDF2Excel>pdf2excel
Investintech.com Inc. PDF2Excel - Command Line Interface
Copyright 2007 Investintech.com Inc. All rights reserved

Unlimited edition

Usage:
  PDF2Excel -c<convert> [-s<schedule>] [-t<template>] [-p<range>] -i<input>
  > -o<output>

  where:
  <convert> specifies type of conversion operation
  <schedule> is time at which conversion will occur
  <template> is name of template file describing output file format
  <range> indicates page range
  <input> specifies input pdf document (may include wildcards)
  <output> specifies target output file path and extension

Note: input and output files must contain full type extension
Wild cards:
  ? Matches any single character in document name
  * Matches 0 or more characters in document name
Conversion types:
  csv Convert PDF to Comma Separated Values
  xls Convert PDF to Excel (XLS format)

Example:
  PDF2Excel -cxls -iSource.pdf -oTarget.xls
  PDF2Excel -ccsv -iSource.pdf -oTarget.txt
  PDF2Excel -ccsv -i*.pdf -o*.txt
  PDF2Excel -cxls -s15:00:00 -iSource.pdf -oTarget.xls
  PDF2Excel -cxls -p5-9,15 -iSource.pdf -oTarget.xls
  PDF2Excel -cxls -p5-9,15o -iSource.pdf -oTarget.xls
  PDF2Excel -cxls -p4-12,16 -iSource.pdf -oTarget.xls
  PDF2Excel -cxls -p4-12,16e -iSource.pdf -oTarget.xls
  PDF2Excel -cxls -tTemplate.t2ae -iSource.pdf -oTarget.xls

  -h or -help This help

C:\Program Files\Investintech\PDF2Excel>
```

The help message describes command usage, some examples and possible conversion types. The command line format is of the form:

```
PDF2Excel -c<conversion_type> [-s<schedule>] [-t<template>]
[-p<range>] -i<input_file(s)> -o<output_file((s))>
```

Conversion type

This parameter determines source and destination file formats. Source file is the file you want to convert. Destination file is the file in which results of the conversion will be put into. It is

PDF2Excel Command Line

important to know that the source file will remain intact. Possible source file formats are Portable Document Format (PDF) and plain text. Possible destination file formats for PDF source are Microsoft Excel (XLS), Comma Separated Values (CSV).

Following table shows the possible values of the `-conversion_type` parameter:

<i>Conversion type</i>	<i>Description</i>
-xls	Converts PDF file to Excel 2.1 file so any version of Microsoft Excel can open destination file.
-csv	Converts PDF file to CSV file so virtually all spreadsheet programs can open destination file.

Schedule

This optional argument allows you to automate a one-off or recurring conversion operation at some time in the future.

A valid PDF-to-Excel schedule parameter excludes whitespace and is of the form:

`time[/EVERY:date[,...]/NEXT:date[,...]]`

where time is a colon delimited 24 hour time to the nearest second e.g. hh:mm:ss

hour: 00 - 23

minute: 00 - 59

second 00 - 59

date either a day of the week e.g. M,T,W,Th,F,S,Su

or a day of the month e.g. 1,2,3,...31

/EVERY clause, /NEXT clause are optional but only one optional clause is permitted

PDF-TO-EXCEL will construct a valid AT command format e.g.

`AT [\\computername] time [/INTERACTIVE] [/EVERY:date[,...] | /NEXT:date[,...]] "command"`

Template

This optional argument allows you to use a custom template providing detailed conversion instructions that assist in the conversion of a PDF Document to an Excel document. Please contact Investintech.com to find out how you can benefit from a custom template.

Range

This optional argument allows use to convert a subset of the PDF document. A continuous range of pages may be specified by the syntax: <starting page>-<ending page> where <starting page> - <ending page> are valid page numbers in the document and the starting page occurs before the ending page.

Example: 10-20

A list of pages may be specified with the syntax: <page i >,<page j>,<page k> where <page i >,<page j>,<page k> are valid page numbers in the document. In addition, you can select even numbered pages or odd numbered pages with the syntax "o" - to select odd numbered pages or "e" - to select even numbered pages.

Input file(s)

This argument determines the source file, i.e. the file you want to convert data from. Please note the following:

- You must include both file name and extension separated with the dot character ('.'). For example, file `business.pdf` consists of file name 'business' and extension 'pdf'.
- You must provide full path to the input file unless it is in the same folder as `PDF-to-Excel.exe`. For example, if input file `business.pdf` is in `MyFiles` folder on C drive, you should type `c:\MyFiles\business.pdf` as an input file argument. If any folder name or file name in the path is longer than 8 characters or contains spaces you must surround its path with double quotation marks, e.g. `"c:\very long path\business.pdf"`

Output file(s)

This argument is a string expression determining the destination file, i.e. the file containing the results of the conversion. The naming rules used by PDF-to-Excel are as follows. PDF-to-Excel will use the specified directory or the current working directory if none is specified. It will use the name corresponding to the name of the source file. It will use the extension Finally Same path and file name rules apply for output file as well as for input file.

Examples of Using PDF-to-Excel

Here are some cases illustrating the use of PDF-to-Excel.

Example #1

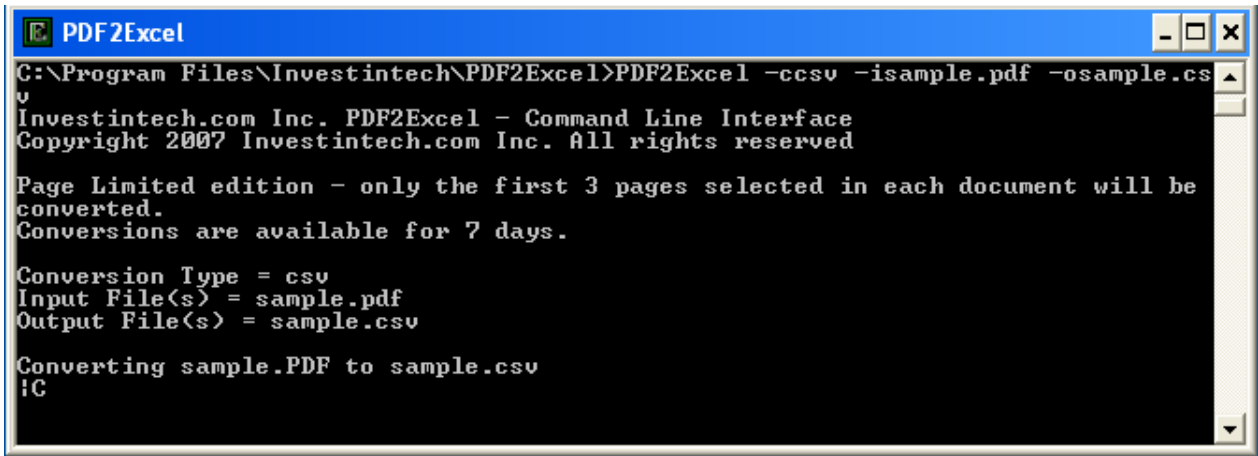
In this example we will assume that both input and output files are in the same folder as PDF-to-Excel. If you type

```
pdf2excel -ccsv sample.pdf sample.csv
```

and press '**Enter**', you will convert from PDF to Comma Separated Value file. The resulting CSV file can be opened by Microsoft Excel. The picture below shows a successful conversion.

PDF2Excel Command Line

For each converted page PDF-to-Excel will print “|” on the screen to let user know that it is active and working.



```
C:\Program Files\Investintech\PDF2Excel>PDF2Excel -ccsv -isample.pdf -osample.csv
Investintech.com Inc. PDF2Excel - Command Line Interface
Copyright 2007 Investintech.com Inc. All rights reserved

Page Limited edition - only the first 3 pages selected in each document will be
converted.
Conversions are available for 7 days.

Conversion Type = csv
Input File(s) = sample.pdf
Output File(s) = sample.csv

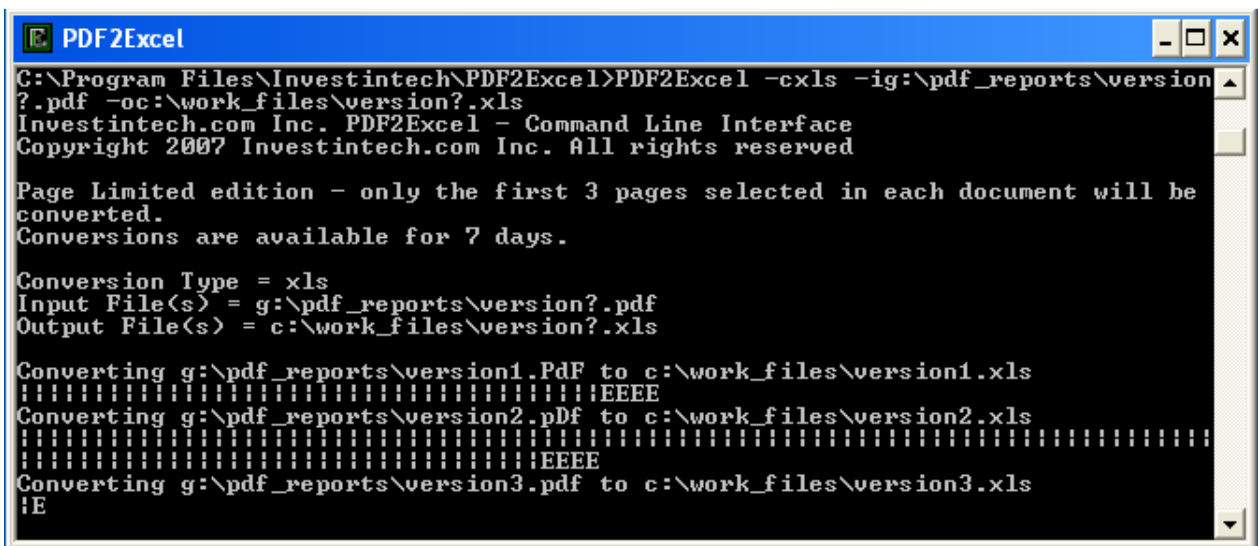
Converting sample.PDF to sample.csv
|C
```

Example #2

In this example we will assume that directory pdf_reports on the G drive contains a series of PDF documents. Further, we assume input file is in the same folder as PDF-to-Excel, but we want output file to be created in the different folder, e.g. d:\conv. Plus, we want the output file format to be Excel workbook format. If you type:

```
PDF2Excel -cxls -ig:\pdf_reports\version?.pdf -
oc:\work_files\version?.xls
```

and press ‘Enter’, you will convert the matching documents from PDF to Excel. The picture below shows a successful conversion.



```
C:\Program Files\Investintech\PDF2Excel>PDF2Excel -cxls -ig:\pdf_reports\version
?.pdf -oc:\work_files\version?.xls
Investintech.com Inc. PDF2Excel - Command Line Interface
Copyright 2007 Investintech.com Inc. All rights reserved

Page Limited edition - only the first 3 pages selected in each document will be
converted.
Conversions are available for 7 days.

Conversion Type = xls
Input File(s) = g:\pdf_reports\version?.pdf
Output File(s) = c:\work_files\version?.xls

Converting g:\pdf_reports\version1.Pdf to c:\work_files\version1.xls
|E
Converting g:\pdf_reports\version2.pdf to c:\work_files\version2.xls
|E
Converting g:\pdf_reports\version3.pdf to c:\work_files\version3.xls
|E
```

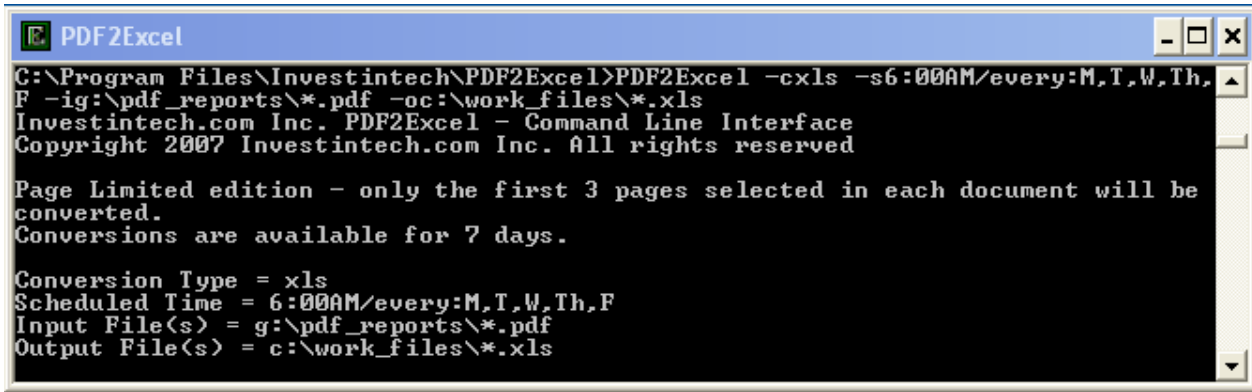
Example #3

In this example we wish to schedule a reoccurring conversion operation. If you type

PDF2Excel Command Line

```
PDF2Excel -cxls -s6:00AM/every:M,T,W,Th,F -ig:\pdf_reports\*.pdf  
-oc:\html_reports\*.xls
```

and press **‘Enter’**, you will schedule the conversion operation. The picture below shows the operation has been scheduled.



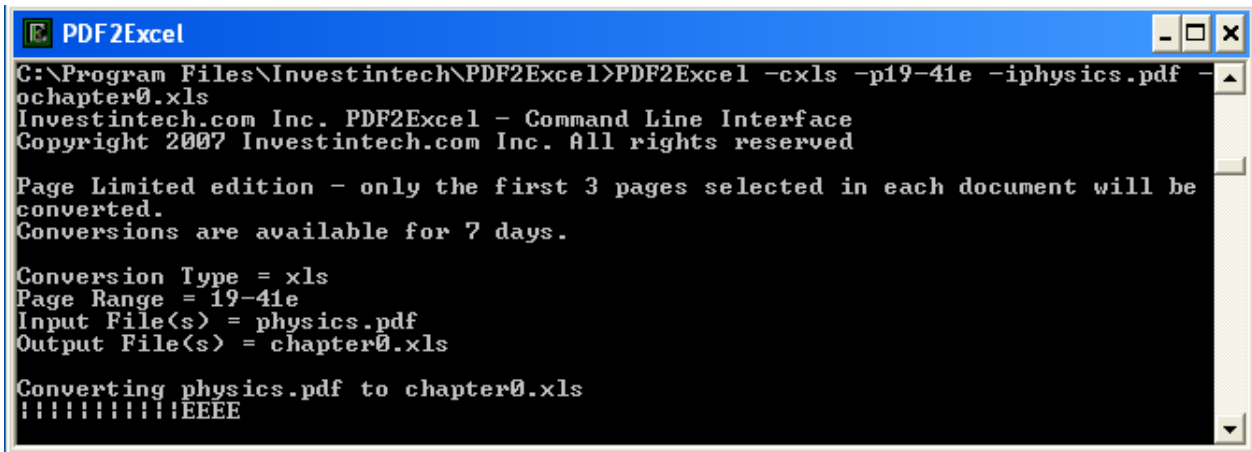
```
PDF2Excel  
C:\Program Files\Investintech\PDF2Excel>PDF2Excel -cxls -s6:00AM/every:M,T,W,Th,  
F -ig:\pdf_reports\*.pdf -oc:\work_files\*.xls  
Investintech.com Inc. PDF2Excel - Command Line Interface  
Copyright 2007 Investintech.com Inc. All rights reserved  
  
Page Limited edition - only the first 3 pages selected in each document will be  
converted.  
Conversions are available for 7 days.  
  
Conversion Type = xls  
Scheduled Time = 6:00AM/every:M,T,W,Th,F  
Input File(s) = g:\pdf_reports\*.pdf  
Output File(s) = c:\work_files\*.xls
```

Example #4

In this example we will convert a range of the document. If you type:

```
PDF2Excel -cxls -p19-41e -iphysics.pdf -ochapter0.xls
```

and press **‘Enter’**, you will convert from PDF to Excel file. The picture below shows a successful conversion.



```
PDF2Excel  
C:\Program Files\Investintech\PDF2Excel>PDF2Excel -cxls -p19-41e -iphysics.pdf -  
ochapter0.xls  
Investintech.com Inc. PDF2Excel - Command Line Interface  
Copyright 2007 Investintech.com Inc. All rights reserved  
  
Page Limited edition - only the first 3 pages selected in each document will be  
converted.  
Conversions are available for 7 days.  
  
Conversion Type = xls  
Page Range = 19-41e  
Input File(s) = physics.pdf  
Output File(s) = chapter0.xls  
  
Converting physics.pdf to chapter0.xls  
!!!!!!!!!!!!EEEE
```

Troubleshooting

There are a few cases in which PDF-to-Excel is unable to perform the requested file conversion. Please consult the following table if you experience any of errors while working with PDF-to-Excel.

<i>Error</i>	<i>Cause</i>	<i>Solution</i>
Unable to convert - did not find input document file name expression.	Misspelled or omitted input document file parameter. It is also possible you do not have permissions to access input file	use -i<filename.pdf> expression to specify PDF document file(s) to be converted or contact your system administrator to give you permissions to read input file.
Unable to convert - did not recognize conversion type:	Misspelled or omitted conversion type parameter	use -cxls or -ccsv command line parameters to specify conversion type
Unable to convert - did not find output document file name expression.xls	Misspelled or omitted input document file parameter. It is also possible you do not have permissions to access input file	Use -o<filename.xls> to specify file(s) that will contain converted information. Try a path for which you are have write permissions.
Empty converted document when using page range	Page range is invalid e.g. -p50-40	Use ascending order in page range e.g. -p40-50
Task not scheduled when using schedule parameter	Schedule time specified incorrectly.	Use a formatted 24 hour time with valid optional recurring sub parameters e.g. -s:30:00 -s06:30:00:/next:1 -23:59:59:/every:M,T,W,Th,F
Input file is damaged or not valid PDF file!	Input file was found but it either is not PDF file or it is damaged or corrupted PDF file	If Acrobat Reader can open the file without any error message please send the file to our technical support e-mail and we will look further into the problem. If Adobe Acrobat (full version) can open the file with error message, save the file back - this way Adobe Acrobat will repair the file. Then try conversion again.

Index

PDF2Excel Command Line, 3, 4
 Conversion type, 12
 Examples of Using, 14
 Input file, 14
 Installation Instructions, 4
 Output file, 13, 14

 System Requirements, 4
 Troubleshooting, 17
 Using, 11
Customer Service, 2
Investintech.com Inc., 1
Technical Support, 2

ERROR: undefined
OFFENDING COMMAND: wnd

STACK:

/Ref_cmd
/Dest
-mark-