



Word Import Tool
User's Guide v1.0



Table of Contents

Terms and Conditions.....	2
System Requirements.....	2
Launch Window.....	3
Import Text.....	4
Define What is Considered a Document.....	4
Result.....	4
Import Tables.....	5
What is imported?.....	5
Tables List.....	6
Data Preview.....	6
Best Guess Data Type.....	7
Setting Data Type or Excluding Columns.....	7
Settings.....	7
Miscellaneous.....	8
Import Images.....	8
Document Summary.....	9

Terms and Conditions

SAS INSTITUTE INC. IS PROVIDING YOU WITH THE COMPUTER SOFTWARE CODE INCLUDED WITH THIS AGREEMENT ("CODE") ON AN "AS IS" BASIS, AND AUTHORIZES YOU TO USE THE CODE SUBJECT TO THE TERMS HEREOF. BY USING THE CODE, YOU AGREE TO THESE TERMS. YOUR USE OF THE CODE IS AT YOUR OWN RISK. SAS INSTITUTE INC. MAKES NO REPRESENTATION OR WARRANTY, EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT AND TITLE, WITH RESPECT TO THE CODE.

The Code is intended to be used solely as part of a product ("Software") you currently have licensed from SAS Institute Inc. or one of its subsidiaries or authorized agents ("SAS"). The Code is designed to either correct an error in the Software or to add functionality to the Software, but has not necessarily been tested. Accordingly, SAS makes no representation or warranty that the Code will operate error-free. SAS is under no obligation to maintain or support the Code.

Neither SAS nor its licensors shall be liable to you or any third party for any general, special, direct, indirect, consequential, incidental or other damages whatsoever arising out of or related to your use or inability to use the Code, even if SAS has been advised of the possibility of such damages.

Except as otherwise provided above, the Code is governed by the same agreement that governs the Software. If you do not have an existing agreement with SAS governing the Software, you may not use the Code.

System Requirements

JMP 12 or higher is required to use the Word Import Tool.

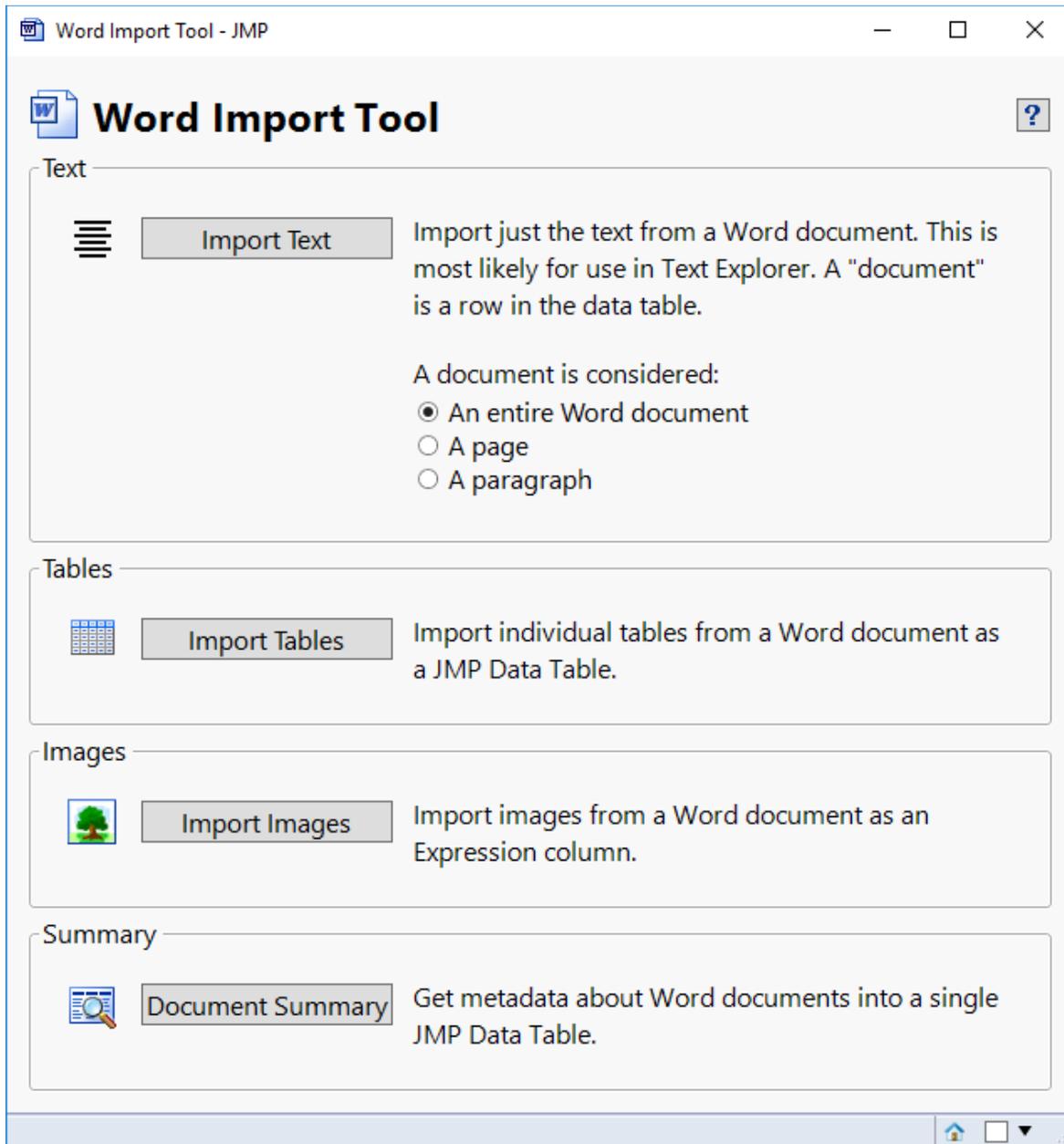
Both Windows and Mac operating systems are supported.

Launch Window

To launch the add-in, click Add-Ins > Word Import Tool



The Word Import Tool's launch window allows you to start the process for importing text, tables, images, or get a summary of your document.



Import Text

NOTE: The term “document” is used to refer to a collection of words within a row in a data table and the term “Word document” refers to a ‘.docx’ file.

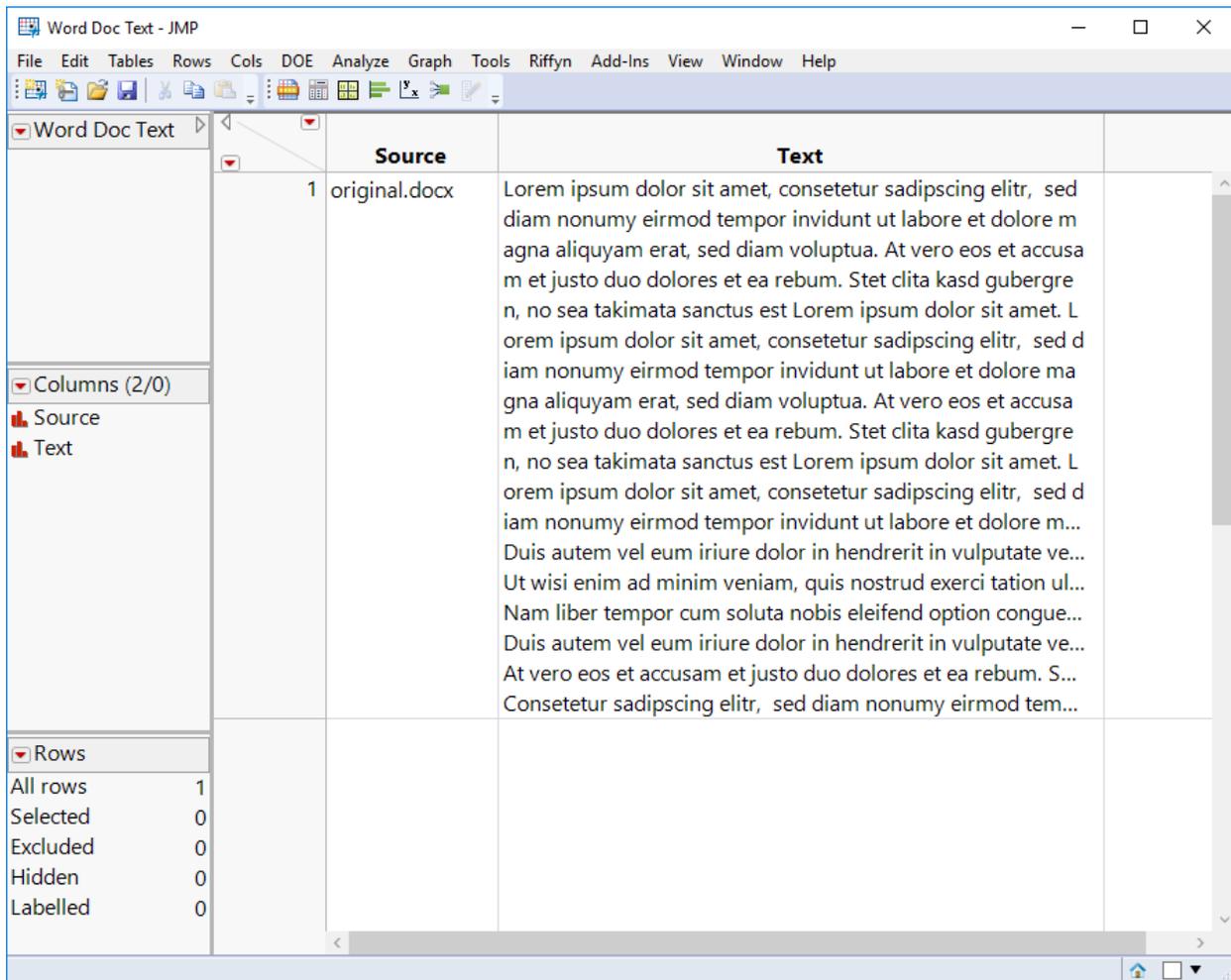
Define What is Considered a Document

To import text from a Word document, first decide what you consider a document. For the Word Import Tool’s Import Text option, you can define a document to be considered an entire Word document’s text, a single page, or a paragraph from the Word document. Text within cells of a table or text boxes will be treated as separate paragraphs.

Choose your document type from the list of options on the launch window before clicking the “Import Text” button.

Result

After selecting your file(s) to be imported, a new data table will be opened with one row per Word document.



The screenshot shows the JMP software interface with a data table titled "Word Doc Text". The table has two columns: "Source" and "Text". The "Source" column contains the value "1 original.docx". The "Text" column contains a large block of Lorem Ipsum placeholder text. The interface also shows a menu bar, a toolbar, and a sidebar with "Columns (2/0)" and "Rows" sections.

	Source	Text
1	original.docx	Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accipit et eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper magna eros ut erat loremIpsum dolor sit amet. Nam liber tempor cum soluta nobis eleifend option congue elit ut erat ea elita, sed ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accipit et eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. S... Consetetur sadipscing elitr, sed diam nonumy eirmod tem...

Import Tables

What is Imported?

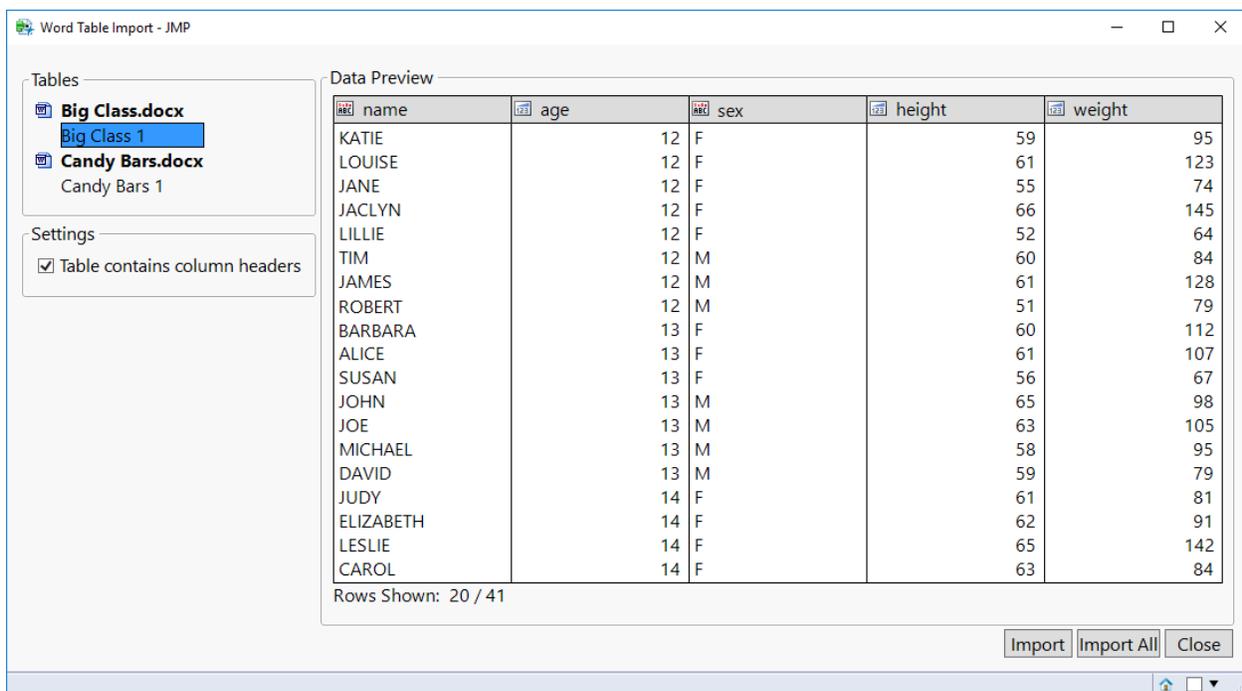
The Import Tables option will look for tables embedded in your Word document(s) and allows you to import each of them as a JMP Data Table.

For example, below is a snippet from what is contained within an example Word document.

Big Class

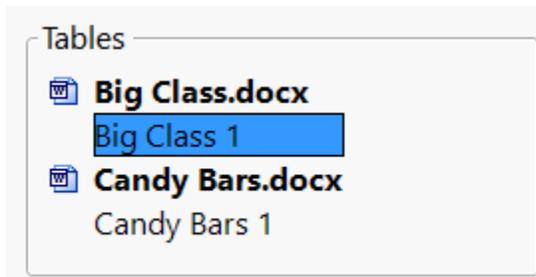
name	age	sex	height	weight
KATIE	12	F	59	95
LOUISE	12	F	61	123
JANE	12	F	55	74
JACLYN	12	F	66	145

After all tables have been found, a preview window is shown with best guess options applied.



Tables List

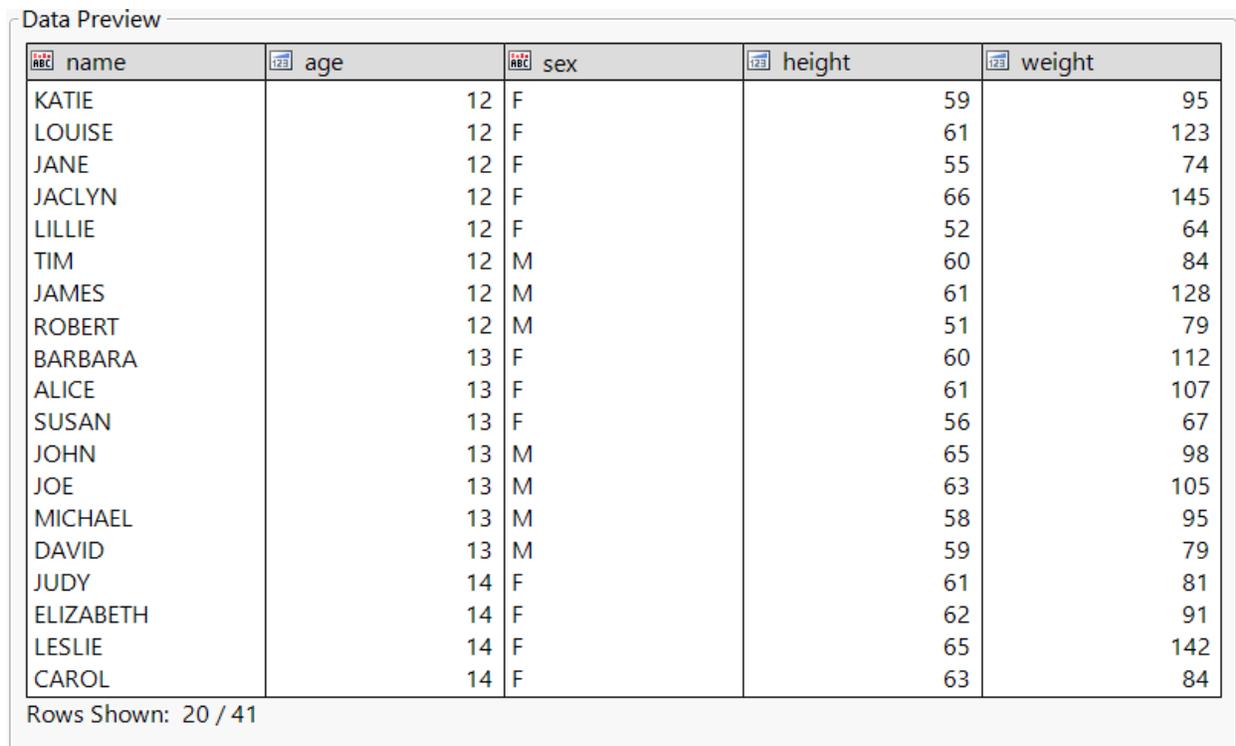
On the left, you can select a table to preview. These tables are grouped by Word document and ordered they appear in the document.



When a selected table has focus (shown by a black outline), you can use the UP and DOWN arrow keys to cycle through each of the tables.

Data Preview

Below is what an imported table will look like in the data preview mode. The first 20 rows of the data will be displayed along with the column names and data type.



The screenshot shows a 'Data Preview' pane containing a table with 5 columns: name, age, sex, height, and weight. The table contains 20 rows of data. Below the table, it says 'Rows Shown: 20 / 41'.

name	age	sex	height	weight
KATIE	12	F	59	95
LOUISE	12	F	61	123
JANE	12	F	55	74
JACLYN	12	F	66	145
LILLIE	12	F	52	64
TIM	12	M	60	84
JAMES	12	M	61	128
ROBERT	12	M	51	79
BARBARA	13	F	60	112
ALICE	13	F	61	107
SUSAN	13	F	56	67
JOHN	13	M	65	98
JOE	13	M	63	105
MICHAEL	13	M	58	95
DAVID	13	M	59	79
JUDY	14	F	61	81
ELIZABETH	14	F	62	91
LESLIE	14	F	65	142
CAROL	14	F	63	84

Rows Shown: 20 / 41

Best Guess Data Type

If the first row contains all numeric values or any character string longer than 50 characters, the add-in will guess that the table does not have column headers. Otherwise, it is guessed that column headers are in row 1.

Below is a table showing how data types are determined for each table.

	First Row is Character	First Row is Numeric
Second Row is Character	Character	Character
Second Row is Numeric	Numeric	Numeric

Setting Data Type or Excluding Columns

If the add-in's best guess did not get the data type correct, you can change it by clicking the icon to the left of the column name. Clicking this icon will cycle through the data type options (and an excluded option). Excluding a column will cause it to not be in the imported table.

Please note that any characters located within a numeric column will not appear as missing until imported as a data table.

Below is a screenshot showing the different icons you can click to change the data type, or exclude a column.

Character	Numeric	Excluded
abc name	123 age	no sex
KATIE	12	F
LOUISE	12	F
JANE	12	F
JACLYN	12	F
LILLIE	12	F

Settings

You can specify whether a table contains column headers or not in row 1.

Settings

Table contains column headers

Miscellaneous

There are two items to know about how certain scenarios are handled.

- Values in merged cells are repeated across all merged cells.
- If a table is located within another table, the nested table cannot be imported and the text within the table will be included in the top parent’s cell value.

Import Images

To import images from a Word document, click the “Import Images” button on the launch window. You can select multiple files or just a single Word document.

Below is the table that is generated by the add-in when importing images from this User’s Guide document.

Source	Image	Page	Name	Title	Description										
1 Word Import Tool User's Guide.docx		1	Group 35												
2 Word Import Tool User's Guide.docx		1	Group 35												
3 Word Import Tool User's Guide.docx		2	Picture 1												
4 Word Import Tool User's Guide.docx		3	Picture 3												
5 Word Import Tool User's Guide.docx		4	Picture 13												
6 Word Import Tool User's Guide.docx	<table border="1"> <thead> <tr> <th>name</th> <th>age</th> <th>sex</th> <th>height</th> <th>weight</th> </tr> </thead> <tbody> <tr> <td>KATIE</td> <td>12</td> <td>F</td> <td>50</td> <td>95</td> </tr> </tbody> </table>	name	age	sex	height	weight	KATIE	12	F	50	95	4	Picture 235		
name	age	sex	height	weight											
KATIE	12	F	50	95											

Note: The ‘Title’ and ‘Description’ columns are grabbed from an image’s Alt Text properties within Microsoft Word. Below is an example table that included an image with Alt Text.

Source	Image	Page	Name	Title	Description
Image with title and description.do...		1	Picture 1	Help Icon	This is a Help Icon

Document Summary

To get a summary of a group of Word documents, click the “Document Summary” button on add-in’s the launch window.

The summary table includes various pieces of information about each document. This information comes from OS level file properties as well as Microsoft Word specific properties.

Below is a list of all columns created in the output data table.

- Filename
- FileSize
- CreateDate
- ModDate
- Creator
- LastModifiedBy
- PageCount
- WordCount
- CharacterCount
- CharacterCountWithSpaces
- LineCount
- ParagraphCount
- TableCount
- ImageCount
- Application
- Template
- Company
- Title
- Subject
- Keywords
- Description
- Revision
- LastPrinted

Below is a portion of a document summary table for two Word documents.

Filename	FileSize	CreateDate	ModDate	Creator	LastModifiedBy	PageCount	WordCount	CharacterCount	CharacterCount WithSpaces	LineCount	ParagraphCount
Big Class.docx	13764	2017-06-26T...	2017-06-26T...	Justin Chilton	Justin Chilton	2	119	680	798	5	1
Candy Bars....	23415	2017-06-26T...	2017-06-26T...	Justin Chilton	Justin Chilton	2	721	4115	4827	34	9