

---

# IBM InfoSphere Information Server V11.3

Writing data to a Microsoft Excel spreadsheet in  
IBM InfoSphere DataStage

© 2014 IBM Corporation



This presentation discusses how to write data to a Microsoft® Excel spreadsheet using the Unstructured Data Stage in IBM InfoSphere® DataStage® version 11.3.

## Objectives

- Create a DataStage job
- Configure Sequential File Stage
- Configure Unstructured Data Stage
- Compile and run job
- View output of job

The objectives of this presentation are to show how to create a DataStage job that reads from a Sequential File stage and writes to a Microsoft Excel spreadsheet using the Unstructured Data stage. This presentation includes step-by-step instructions starting from creation and configuration of the job to running the job and viewing the target data.

# Writing data to a Microsoft Excel spreadsheet

- Input sequential file – CSV format
  - Employee.txt – employee data

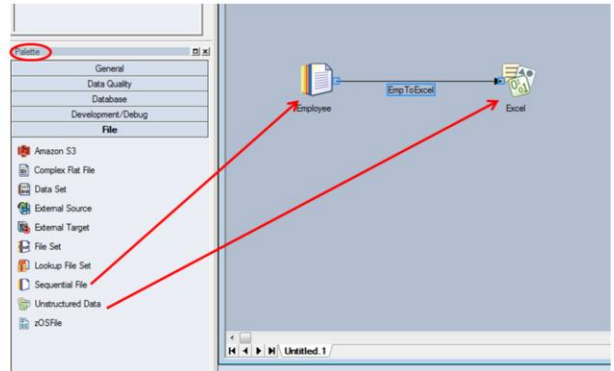
```

Employee - Notepad
File Edit Format View Help
"000010","CHRISTINE","I","HAAS","ANN","3978","1995-01-01","PRES","18","F","1963-08-24","0152750.00","0001000.00","0004220.00"
"000020","MICHAEL","L","TEMPERSON","301","3476","2003-10-10","MANAGER","18","M","1978-02-02","0094350.00","0000000.00","0001300.00"
"000030","SALLY","A","OWAN","C01","4738","2005-04-05","MANAGER","20","F","1971-05-11","0096250.00","0000000.00","0003060.00"
"000040","JOHN","B","GEYER","E01","6789","1979-08-17","MANAGER","16","M","1955-09-15","0008175.00","0000000.00","0003214.00"
"000050","IRVING","F","STERN","D11","6433","2003-09-14","MANAGER","16","M","1975-07-07","0072250.00","0000500.00","0002500.00"
"000060","EVA","D","PULASKI","D21","7831","2005-09-30","MANAGER","16","F","2003-05-26","0096170.00","0000700.00","0002893.00"
"000070","EILEEN","W","HENDERSON","E11","5498","2000-08-15","MANAGER","16","F","1971-05-15","0009750.00","0000600.00","0002380.00"
"000080","THEODORE","G","SPENCER","E21","0972","2000-06-19","MANAGER","14","M","1980-12-18","0005150.00","0000500.00","0002092.00"
"000090","VICKERZO","G","LUICCHESI","A00","3490","1988-05-16","SALESREP","19","M","1959-11-05","0066500.00","0000900.00","0003720.00"
"000100","SEAN","J","CORRELL","ANN","2167","1993-12-05","CLERK","14","M","1972-10-18","0049250.00","0000600.00","0002340.00"
"000110","DELORES","M","QUINTANA","C01","4578","2001-07-20","ANALYST","16","F","1955-09-15","0073000.00","0000500.00","0001900.00"
"000120","HEATHER","A","NICHOLLS","C01","1793","2006-12-15","ANALYST","18","F","1976-01-19","0068420.00","0000600.00","0002274.00"
"000130","BRUCE","J","ADAMSON","D11","4510","2002-02-12","DESIGNER","16","M","1977-05-17","0052800.00","0000500.00","0002022.00"
"000140","ELIZABETH","M","PEAKAR","D11","3782","2006-10-11","DESIGNER","17","F","1980-04-12","0062250.00","0000400.00","0001700.00"
"000150","MASATOSHI","J","YOSHIMURA","D11","2898","1999-09-15","DESIGNER","16","M","1981-01-05","0044600.00","0000500.00","0001974.00"
"000160","MARILYN","S","SCOTTEN","D11","1482","2003-07-07","DESIGNER","17","F","1979-02-21","0051140.00","0000500.00","0001707.00"
"000170","JAMES","M","MILNER","D11","2960","2004-07-16","DESIGNER","16","M","1982-06-25","0050450.00","0000000.00","0001636.00"
"000180","DAVID","J","BRONK","D11","4501","2002-03-03","DESIGNER","16","M","1971-05-29","0057740.00","0000600.00","0002217.00"
"000190","WILLIAM","I","JONES","D11","0942","1998-04-11","DESIGNER","17","M","2003-02-23","0068270.00","0000400.00","0001462.00"
"000200","JENNIFER","K","LITZ","D11","0072","1998-08-29","DESIGNER","18","F","1978-03-19","0049840.00","0000600.00","0001387.00"
"000210","JAMES","J","JEFFERSON","D21","2094","1996-11-21","CLERK","14","M","1980-05-30","0042180.00","0000400.00","0001774.00"
"000220","SALVATORE","M","MARINO","D21","3780","2004-12-05","CLERK","17","M","2002-03-31","0048760.00","0000600.00","0002301.00"
"000230","DANIEL","S","SMITH","D21","0961","1999-10-30","CLERK","15","M","1969-11-12","0049180.00","0000000.00","0001554.00"
"000240","DAVID","F","JOHNSON","D21","8953","2005-09-11","CLERK","16","F","1976-10-05","0047250.00","0000300.00","0001300.00"
"000250","MARIA","L","PERE","D21","9001","2006-09-30","CLERK","15","F","2003-05-26","0037380.00","0000500.00","0002100.00"
"000260","ETHEL","M","SCHEIDER","E11","8997","1997-03-24","OPERATOR","17","F","1976-03-20","0036200.00","0000500.00","0002100.00"
"000270","JOHN","R","PARKER","E11","4502","2006-05-30","OPERATOR","12","M","1985-07-09","0035340.00","0000300.00","0001227.00"
"000280","PHILIP","K","SMITH","E11","2095","2002-06-19","OPERATOR","14","M","1976-10-27","0037750.00","0000400.00","0001420.00"
"000290","RALPH","F","SEIBERT","E11","3332","1994-00-12","OPERATOR","12","F","1961-04-21","0035900.00","0000300.00","0001272.00"
"000300","RAMLA","V","MENTA","E21","9990","1995-07-07","FIELDREP","16","M","1962-08-11","0039950.00","0000400.00","0001596.00"
"000310","WING","L","LEE","E21","2183","2006-02-23","FIELDREP","14","M","1971-07-18","0045370.00","0000500.00","0002030.00"
"000320","JASON","M","GORDY","E21","5690","1977-05-05","FIELDREP","16","M","1956-05-17","0041840.00","0000500.00","0001907.00"
"000330","DIAN","J","HOPFINGER","ANN","3978","1995-01-01","SALESREP","18","F","1973-08-14","0046500.00","0001000.00","0004220.00"
    
```

This presentation uses a text file, Employee.txt, as the source data. The source file contains information of employees in CSV, comma-separated value, format. The slides to follow demonstrate how this information can be written to a Microsoft Excel spreadsheet.

## Creating the DataStage job

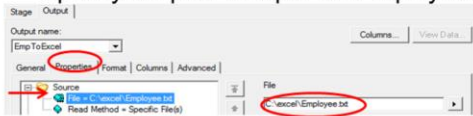
- Open Designer
  - New => Parallel job
- File section of palette
  - Add Sequential File stage
  - Add Unstructured Data stage
- Link Sequential File stage to Unstructured Data stage
- Rename stages and links



The first step is to create a new DataStage parallel job. Open the DataStage designer client and pick New => Parallel job. From the File section of the Palette, drag a Sequential File stage and an Unstructured Data stage to the canvas and link the two together. To make the job design on the canvas more readable, right-click each stage and the link and rename them appropriately.

## Configuring the Sequential File stage (1 of 2)

- Double-click Sequential File stage
- Properties tab
  - Specify file path of input file Employee.txt



- Columns tab
  - Define columns

Column name	Key	SQL type	Extended	Length	Scale	Nullable
1 EMPNO	<input checked="" type="checkbox"/>	Char			6	No
2 FIRSTNAME	<input type="checkbox"/>	VarChar		12		No
3 MIDINIT	<input type="checkbox"/>	Char		1		Yes
4 LASTNAME	<input type="checkbox"/>	VarChar		15		No
5 WORKSHEPT	<input type="checkbox"/>	Char		3		Yes
6 PHONENO	<input type="checkbox"/>	Char		4		Yes
7 HIREDATE	<input type="checkbox"/>	Date		10		Yes
8 JOB	<input type="checkbox"/>	Char		8		Yes
9 EDLEVEL	<input type="checkbox"/>	Smallint		2		No
10 SEX	<input type="checkbox"/>	Char		1		Yes
11 BIRTHDATE	<input type="checkbox"/>	Date		10		Yes
12 SALARY	<input type="checkbox"/>	Decimal		9	2	Yes
13 BONUS	<input type="checkbox"/>	Decimal		9	2	Yes
14 COMM	<input type="checkbox"/>	Decimal		9	2	Yes

5

© 2014 IBM Corporation

In this example, the Sequential File stage is used as the source. Double-click the Sequential File stage to edit the properties. On the Properties tab, specify the path to the input file Employee.txt. Next, on the Columns tab, define the columns and click OK to save the changes.

## Configuring the Sequential File stage (2 of 2)

- Right-click Sequential File Stage => View Employee Data

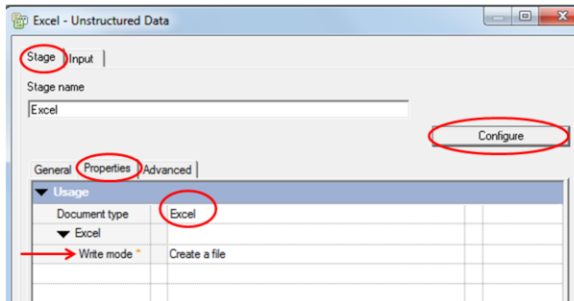
WriteToExcel..Employee.EmpToExcel - Data Browser

EMPNO	FIRSTNAME	MIDINIT	LASTNAME	WORKDEPT	PHONENO	HIREDATE	JOB	EDLEVEL	SEX	BIRTHDATE	SALARY	BONUS	COMM
000010	CHRISTINE	I	HAAS	A00	3978	1995-01-01	PRES	18	F	1963-08-24	0152750.0	0001000.0	0004220.0
000020	MICHAEL	L	THOMPSON	B01	3476	2003-10-10	MANAGER	18	M	1978-02-02	0094250.0	0000800.0	0003300.0
000030	SALLY	A	KWAN	C01	4738	2005-04-05	MANAGER	20	F	1971-05-11	0098250.0	0000800.0	0003060.0
000050	JOHN	B	GEYER	E01	6789	1979-08-17	MANAGER	16	M	1955-09-15	0080175.0	0000800.0	0003214.0
000060	IRVING	F	STERN	D11	6423	2003-09-14	MANAGER	16	M	1975-07-07	0072250.0	0000500.0	0002580.0
000070	EVA	D	PULASKI	D21	7831	2005-09-30	MANAGER	16	F	2003-05-26	0096170.0	0000700.0	0002895.0
000090	EILEEN	W	HENDERSON	E11	5498	2000-08-15	MANAGER	16	F	1971-05-15	0089750.0	0000600.0	0002380.0
000100	THEODORE	Q	SPENSER	E21	0972	2000-06-19	MANAGER	14	M	1980-12-18	0086150.0	0000500.0	0002092.0
000110	VINCENZO	G	LUCCHESSI	A00	3490	1988-05-16	SALESREP	19	M	1959-11-05	0066500.0	0000900.0	0003720.0
000120	SEAN		O'CONNELL	A00	2167	1993-12-05	CLERK	14	M	1972-10-18	0049250.0	0000600.0	0002340.0
000130	DELORES	M	QUINTANA	C01	4578	2001-07-28	ANALYST	16	F	1955-09-15	0073800.0	0000500.0	0001904.0
000140	HEATHER	A	NICHOLLS	C01	1793	2006-12-15	ANALYST	18	F	1976-01-19	0068420.0	0000600.0	0002274.0
000150	BRUCE		ADAMSON	D11	4510	2002-02-12	DESIGNER	16	M	1977-05-17	0055280.0	0000500.0	0002022.0
000160	ELIZABETH	R	PIANKA	D11	3782	2006-10-11	DESIGNER	17	F	1980-04-12	0062250.0	0000400.0	0001780.0
000170	MASATOSHI	J	YOSHIMURA	D11	2890	1999-09-15	DESIGNER	16	M	1981-01-05	0044680.0	0000500.0	0001974.0
000180	MARILYN	S	SCOUTTEN	D11	1682	2003-07-07	DESIGNER	17	F	1979-02-21	0051340.0	0000500.0	0001707.0
000190	JAMES	H	WALKER	D11	2986	2004-07-26	DESIGNER	16	M	1982-06-25	0050450.0	0000400.0	0001636.0
000200	DAVID		BROWN	D11	4501	2002-03-03	DESIGNER	16	M	1971-05-29	0057740.0	0000600.0	0002217.0
000210	WILLIAM	T	JONES	D11	0942	1998-04-11	DESIGNER	17	M	2003-02-23	0068270.0	0000400.0	0001462.0
000220	JENNIFER	K	LUTZ	D11	0672	1998-08-29	DESIGNER	18	F	1978-03-19	0049840.0	0000600.0	0002387.0
000230	JAMES	J	JEFFERSON	D21	2094	1996-11-21	CLERK	14	M	1980-05-30	0042180.0	0000400.0	0001774.0
000240	SALVATORE	M	MARINO	D21	3780	2004-12-05	CLERK	17	M	2002-03-31	0048760.0	0000600.0	0002301.0

Next, right-click the Sequential File stage and select “View Employee Data” to view the Employee.txt file’s content. This step validates that the Sequential File stage is configured properly.

## Configuring the Unstructured Data stage (1 of 2)

- Double-click Unstructured Data
- Stage tab => Document type => Excel
- Write mode = Create a file
- Click Configure



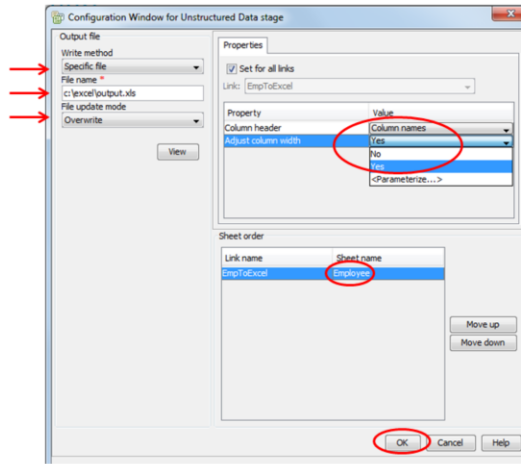
7

© 2014 IBM Corporation

Next, double-click the Unstructured Data stage to open the stage properties. On the Stage tab, select Excel from the Document type list and select Create a file from the Write mode list. Click Configure.

## Configuring the Unstructured Data stage (2 of 2)

- Write method = Specific file
- File name = Full path to output file
- File update mode = Overwrite
- Property section
  - Column header = Column names
  - Adjust column width = Yes
- Sheet order section
  - Sheet name = Employee
- Click OK
- Click OK on the stage properties

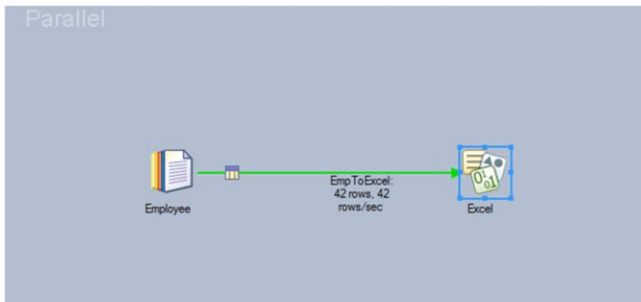


In the configuration window that opens after clicking Configure, select “Specific file” for the Write method, enter the full path to the output file including the file name, and choose Overwrite for the file update mode. Next, under the Property section, set Column header to Column names if you want the first row to be the column names. Set it to None if the first row should be data. Set the Adjust column width to Yes to adjust the column width in the sheet to fit the contents of the column. Finally, enter in the sheet name and click OK. This brings you back to the main property page for the stage. Click OK to save.



## Compiling and running the job

- Save job
- Compile and run job



Next, save and compile the job. Run the job and view the output.

## Viewing the output of the job

- Open Output.xls to view results

EMPNO	FIRSTNAME	MIDINIT	LASTNAME	WORKDEPT	PHONENO	HIREDATE	JOB	EDLEVEL	SEX	BIRTHDATE	SALARY	BONUS	COMM
000010	CHRISTINE	I	HAAS	A00	3978	1995-01-01	PRES		18 F	1963-08-24	152750	1000	4220
000020	MICHAEL	L	THOMPSON	B01	3476	2003-10-10	MANAGER		18 M	1978-02-02	94250	800	3300
000030	SALLY	A	KWAN	C01	4738	2005-04-05	MANAGER		20 F	1971-05-11	98250	800	3060
000050	JOHN	B	GEYER	E01	6789	1979-08-17	MANAGER		16 M	1955-09-15	80175	800	3214
000060	IRVING	F	STERN	D11	6423	2003-09-14	MANAGER		16 M	1975-07-07	72250	500	2580
000070	EVA	D	PULASKI	D21	7831	2005-09-30	MANAGER		16 F	2003-05-26	96170	700	2893
000090	EILEEN	W	HENDERSON	E11	5498	2000-08-15	MANAGER		16 F	1971-05-15	89750	600	2380
000100	THEODORE	Q	SPENSER	E21	0972	2000-06-19	MANAGER		14 M	1980-12-18	86150	500	2092
000110	VINCENZO	G	LUCCHESSI	A00	3490	1988-05-16	SALESREP		19 M	1959-11-05	66500	900	3720
000120	SEAN		O'CONNELL	A00	2167	1993-12-05	CLERK		14 M	1972-10-18	49250	600	2340
000130	DELORES	M	QUINTANA	C01	4578	2001-07-28	ANALYST		16 F	1955-09-15	73800	500	1904
000140	HEATHER	A	NICHOLLS	C01	1793	2006-12-15	ANALYST		18 F	1976-01-19	68420	600	2274
000150	BRUCE		ADAMSON	D11	4510	2002-02-12	DESIGNER		16 M	1977-05-17	55280	500	2022
000160	ELIZABETH	R	PIANKA	D11	3782	2006-10-11	DESIGNER		17 F	1980-04-12	62250	400	1780
000170	MASATOSHI	J	YOSHIMURA	D11	2890	1999-09-15	DESIGNER		16 M	1981-01-05	44680	500	1974
000180	MARILYN	S	SCOUTTEN	D11	1682	2003-07-07	DESIGNER		17 F	1979-02-21	51340	500	1707
000190	JAMES	H	WALKER	D11	2986	2004-07-26	DESIGNER		16 M	1982-06-25	50450	400	1636
000200	DAVID		BROWN	D11	4501	2002-03-03	DESIGNER		16 M	1971-05-29	57740	600	2217
000210	WILLIAM	T	JONES	D11	0942	1998-04-11	DESIGNER		17 M	2003-02-23	68270	400	1462

After you run the job, open the file and verify the output. This slide displays the data in Output.xls in Microsoft Excel.

## Trademarks, disclaimer, and copyright information

IBM, the IBM logo, ibm.com, DataStage, and InfoSphere are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of other IBM trademarks is available on the web at "[Copyright and trademark information](http://www.ibm.com/legal/copytrade.shtml)" at <http://www.ibm.com/legal/copytrade.shtml>

Microsoft, and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON IBM'S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS OR SOFTWARE.

© Copyright International Business Machines Corporation 2014. All rights reserved.