

Nisa Hayuningrum
1810120019
UAS Data Warehouse Praktek

Jawaban:

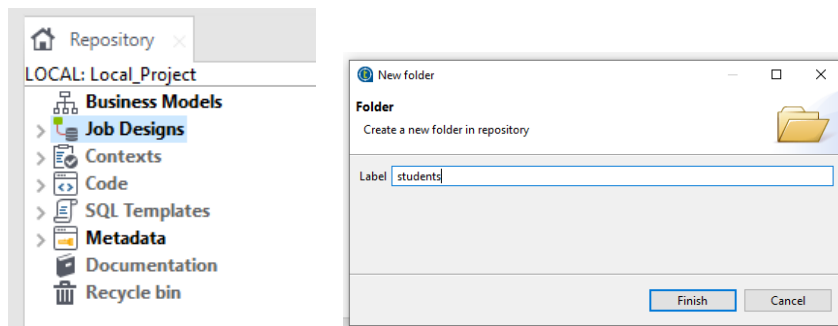
1. Tabel studentcollege.csv

	A	B	C	D	E	F
1	First,Last,Gender,Degree,Major,University					
2	Dina,Amelia,Female,S1,Pendidikan Dokter,UI					
3	Andin,Permata,Female,S1,Akuntansi,UB					
4	Bianca,Afriza,Female,S1,Hubungan Internasional,UGM					
5	Cindy,Syaputri,Female,S1,Sastra Inggris,IPB					
6	Chandra,Kirana,Female,S1,Psikologi,UNY					
7	Devina,Anggraini,Female,S1,Ilmu Komunikasi,UPI					
8	Danilar,Rahmawati,Female,S1,Manajemen,UNDIP					
9	Erhas,Syandika,Male,S1,Ilmu Komputer,UP					
10	Farhan,Febriansyah,Male,S1,Kriminologi,UNES					
11	Gandha,Raharja,Male,S1,Sistem Informasi,UIN					
12	Handika,Putra,Male,S1,Matematika,UNPAD					
13	Intan,Dwirizki,Male,S1,Statistika,UNJ					
14	Johannes,Hutapea,Male,S1,Ilmu Hukum,UKI					
15	Kevin,Sanjaya,Male,S1,Manajemen,UPH					
16	Lukman,Hardian,Male,S1,Biologi,UNHAS					

Tabel University.txt

University - Notepad
File Edit Format View Help
UI, Universitas Indonesia
UII, Universitas Islam Indonesia
IPB, Institut Pertanian Bogor
UB, Universitas Brawijaya
UNJ, Universitas Negeri Jakarta
UKI, Universitas Kristen Indonesia
UNDIP, Universitas Diponegoro
UGM, Universitas Gajah Mada
UNES, Universitas Negeri Semarang
UNY, Universitas Negeri Yogyakarta
UNPAD, Universitas Padjajaran
UPH, Universitas Pelita Harapan
UP, Universitas Pancasila
UPI, Universitas Pendidikan Indonesia Bandung
UNHAS, Universitas Hasanuddin

2. Buka aplikasi Talend -> start now -> klik kanan bagian Job Designs -> create folder dengan nama student.



Klik kanan folder student -> create job -> isi data sebagai berikut -> finish

New Job
Add a job in the repository

Name: lookupmajor

Purpose: convert major to uppercase

Description: uppercase major

Author: user@talend.com

Locker:

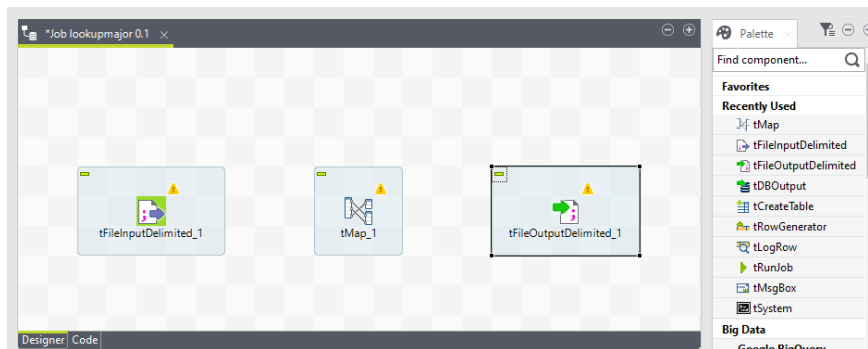
Version: 0.1 ☐ M ☐ m

Status:

Path: students

Lakukan penyetaraan penggunaan huruf besar untuk nama **major**

1. Di bagian palette search lakukan **drag n drop** **tFileinputDelimited**, **tFileoutputDelimited**, dan **tMap** ke worksheet/design.



2. Klik tFileinputDelimited di bagian component ubah pada Basic Settings sebagai berikut:
 - Pilih file input nya yaitu studentcollege.csv
 - ubah Field separator menjadi “,”
 - ubah header menjadi 1.
 - Kemudian edit schema.

Job(lookupmajor 0.1) Contexts(lookupmajor) Component

tFileInputDelimited_1

Basic settings

Property Type: Built-In

Schema: Built-In

"When the input source is a stream or a zip file, footer and random shouldn't be bigger than 0."

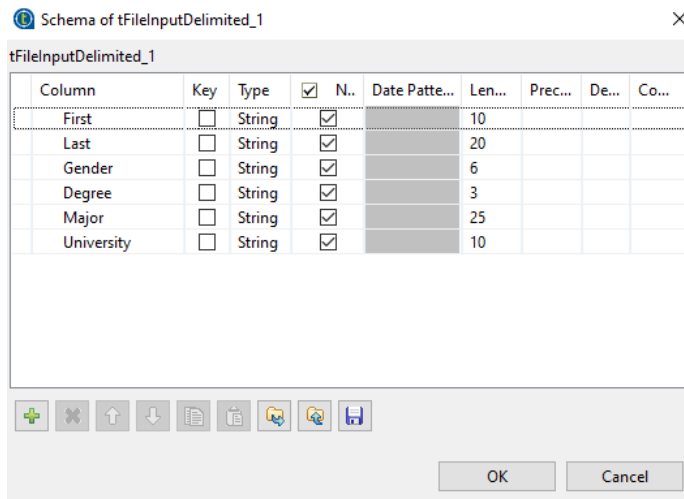
File name/Stream: C:/Users/HP/Documents/Semester 5/DataWarehouse/Tugas/studentcollege.csv

Row Separator: "\n" Field Separator: ","

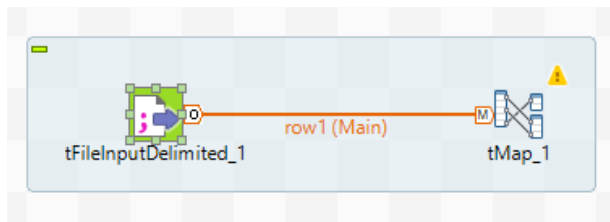
☐ CSV options

Header: 1 Footer: 0 Limit:

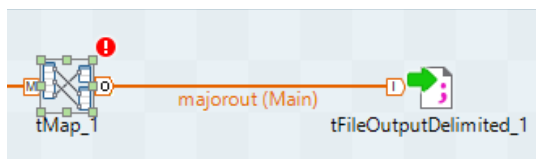
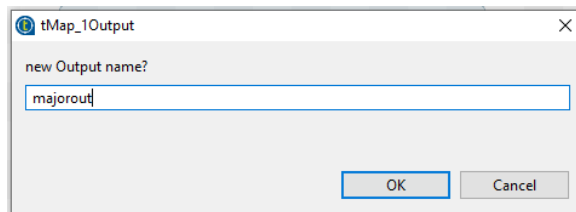
☒ Skip empty rows ☐ Uncompress as zip file ☐ Die on error



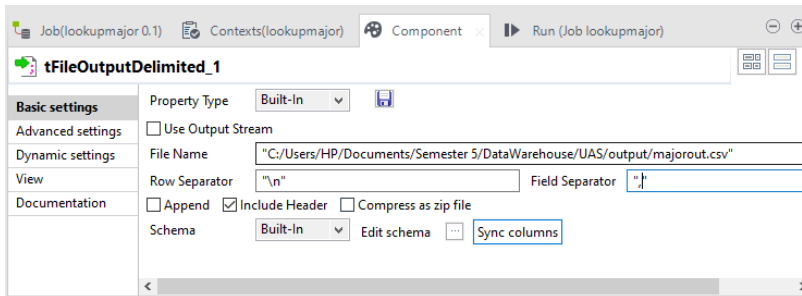
3. Klik kanan **tFileInputDelimited** -> **row** -> **main** -> sambungkan ke **tMap**



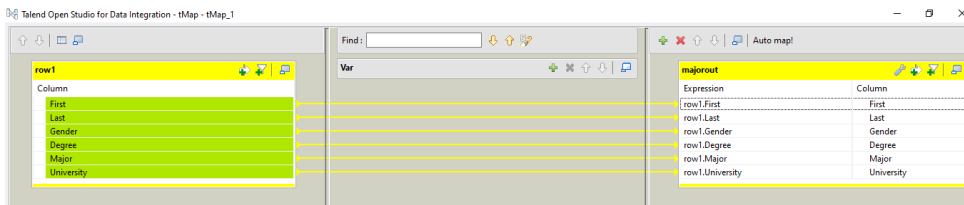
Kemudian klik kanan **tMap** -> **row** -> **new output (main)**, sambungkan ke **tFileOutputDelimited**



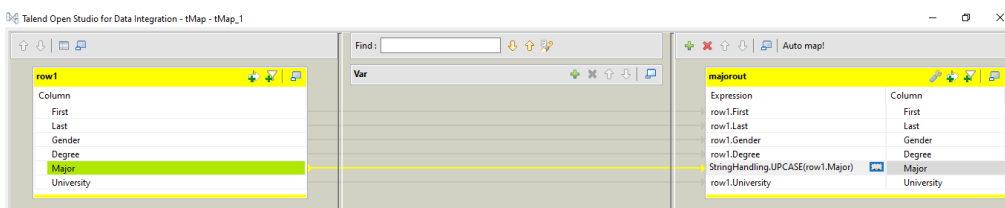
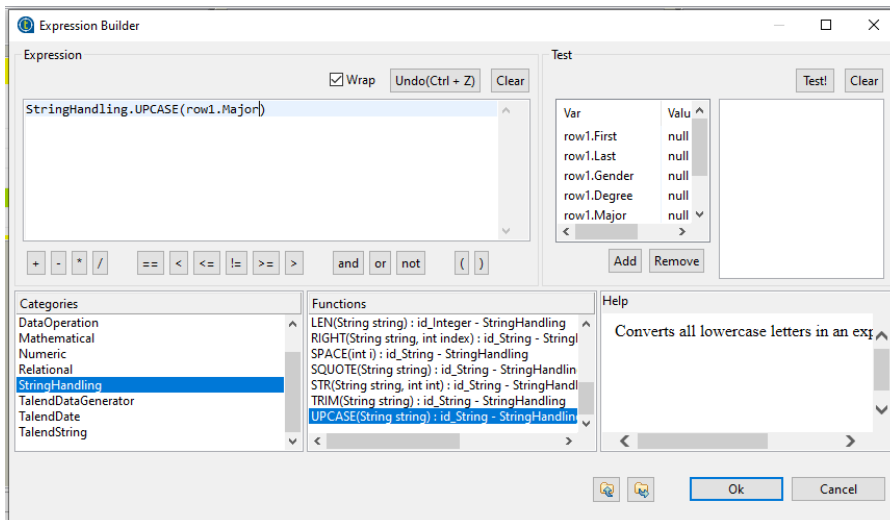
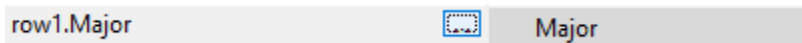
4. Klik **tFileOutputDelimited** dibagian **component** pada **Basic setting** lakukan perubahan sebagai berikut:
- File namanya majorout.csv
 - Field separatornya “,”
 - Centang include header.



5. Kemudian double click **tMap** -> pilih seluruh data dibagian row1 -> drag ke majorout



Klik **row1.Major** di tabel majorout -> **StringHandling** -> **UPCASE(row1.Major)** -> OK



Karena semua sudah terelasi errornya hilang.

3. Kemudian klik **Metadata** dibagian navbar -> klik kanan **File delimited** -> **create file delimited**

New Delimited File

File - Step 1 of 4

Add a Metadata File on repository
Define the properties

Name: UniversityCode

Purpose: memberi nama pada kode universitas

Description: memberi nama di Universitycodes

Author: user@talend.com

Locker:

Version: 0.1

Status:

Path: Select

< Back Next > Finish Cancel

1. Lanjut step 2 -> pilih file **University.txt** -> formatnya **WINDOWS** -> **Next**

New Delimited File

File - Step 2 of 4

Add a Metadata File on repository
Define the path of the file and the format settings

File Settings

Server: localhost 127.0.0.1

File: C:/Users/HP/Documents/Semester 5/DataWarehouse/UAS/University.txt

Format: WINDOWS

File Viewer

- UI, Universitas Indonesia
- UII, Universitas Islam Indonesia
- IPB, Institut Pertanian Bogor
- UB, Universitas Brawijaya
- UNJ, Universitas Negeri Jakarta
- UKI, Universitas Kristen Indonesia
- UNDIP, Universitas Diponegoro
- UGM, Universitas Gajah Mada
- UNES, Universitas Negeri Semarang
- UNY, Universitas Negeri Yogyakarta
- UNPAD, Universitas Padjadjaran
- UPH, Universitas Pelita Harapan
- UP, Universitas Pancasila
- UPI, Universitas Pendidikan Indonesia Bandung

< Next > Finish Cancel

2. Lanjut step 3 -> ubah Field Separator menjadi **Comma** -> **Refresh**

New Delimited File

File - Step 3 of 4
Add a Metadata File on repository
Define the setting of the parse job

File Settings
 Encoding: US-ASCII
 Field Separator: Comma Corresponding Character: ""
 Row Separator: Standard EOL Corresponding Character: "\n"
 Escape Char Settings: ☒ Delimited
 Escape Char: Empty
 Text Enclosure: Empty
☐ Split row before field

Rows To Skip
 If any rows must be ignored, specify the following parameters
 Header: ☐
 Footer: ☐
☐ Skip empty row

Limit Of Rows
 If the number of lines must be limited, specify this number
 Limit: ☐

Preview | **Output**
☐ Set heading row as column names Refresh Preview

Column 0	Column 1
UI	Universitas Indonesia
UII	Universitas Islam Indonesia
IPB	Institut Pertanian Bogor
UB	Universitas Brawijaya

Export as context Revert Context

< Back Next > Finish Cancel

3. Lanjut step 4 -> ubah nama column menjadi **university** dan **university_name**

New Delimited File

File - Step 4 of 4
Add a Schema on repository
Define the Schema

Name: metadata
 Comment:

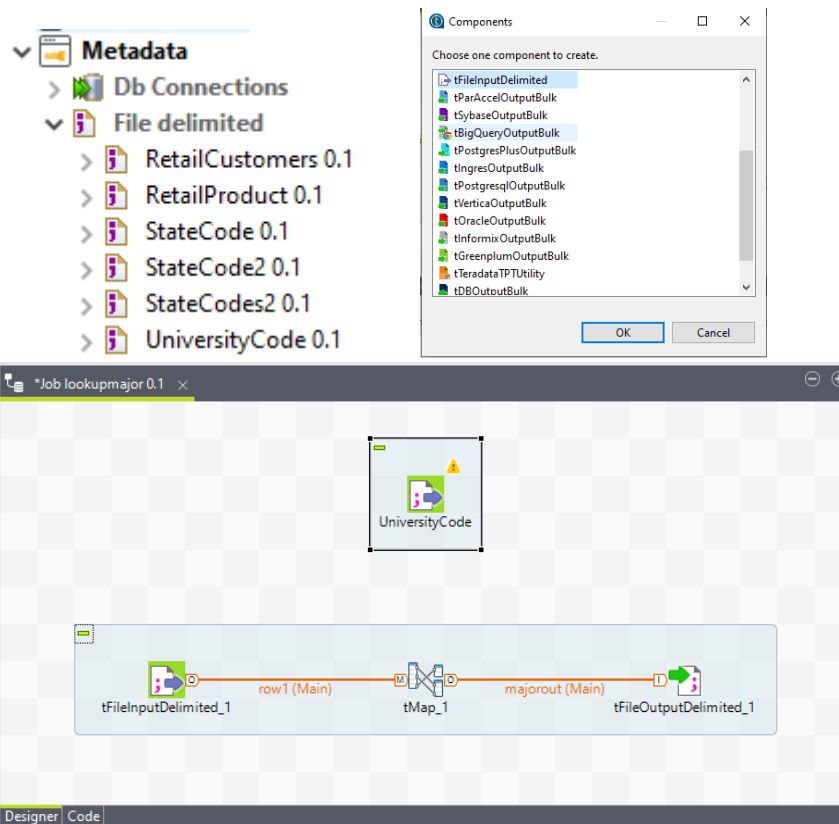
Schema
 Click to update schema preview Guess

Description of the Schema

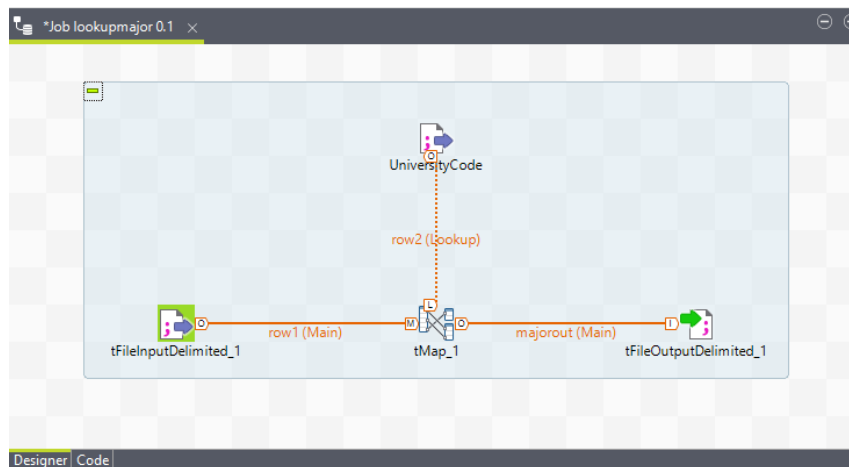
Column	Key	Type	<input checked="" type="checkbox"/> N..	Date Pattern (Ctrl+Sp...	Length	Precision	Default	Comment
university	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		5	0		
university_name	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		25	0		

< Back Next > Finish Cancel

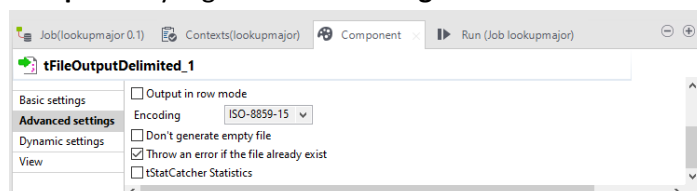
4. Kemudian di File Delimited drag **UniversityCode0.1** ke worksheet/design -> pilih **tFileInputDelimited**



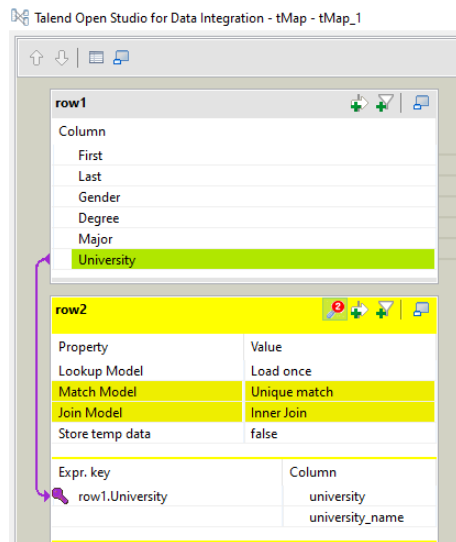
Kemudian lakukan relasi antara UniversityCode dengan tMap -> klik kanan **UniversityCode** -> **Row** -> **Main** -> pilih **tMap**



5. Kemudian di bagian tFileOutputDelimited ceklis **Throw an error if the file already exist** di **Component** yang **Advanced setting**



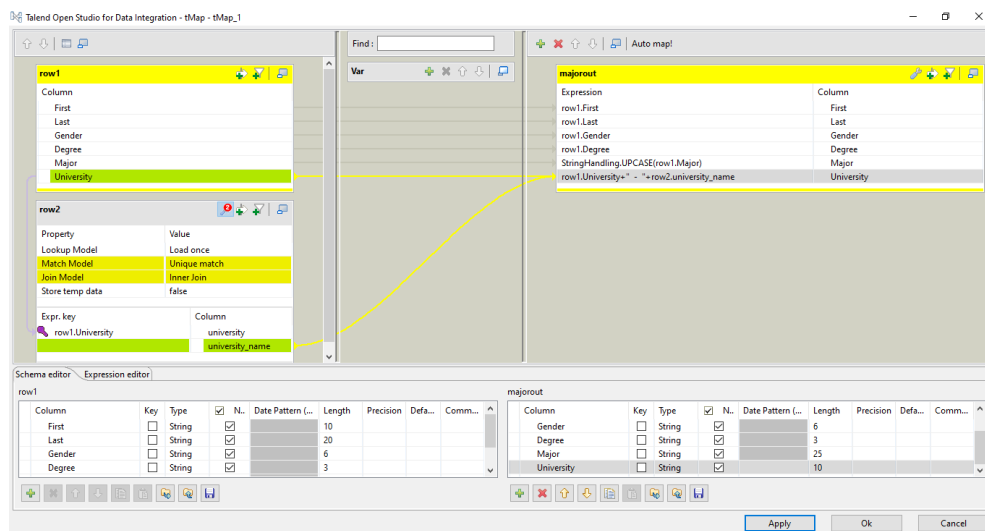
6. Drag **University** di row1 ke **university** di row2 -> klik **setting** di row2->ubah Join Model menjadi **Inner Join**.



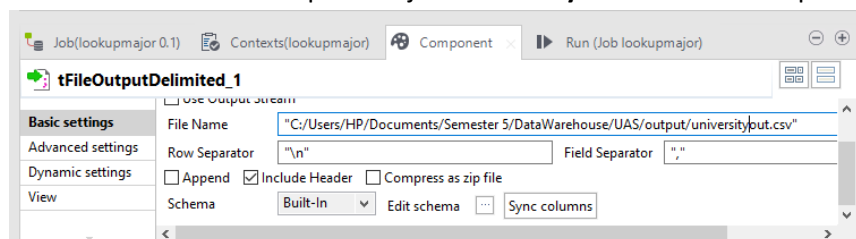
Hubungkan **university_name** di row2 ke **majorout** bagian row1.University. Tambahkan **“ - “** + diantara row1.University dan row2.university_name. Sebagai berikut:

row1.University+“ - “+row2.university_name University

Hasil:

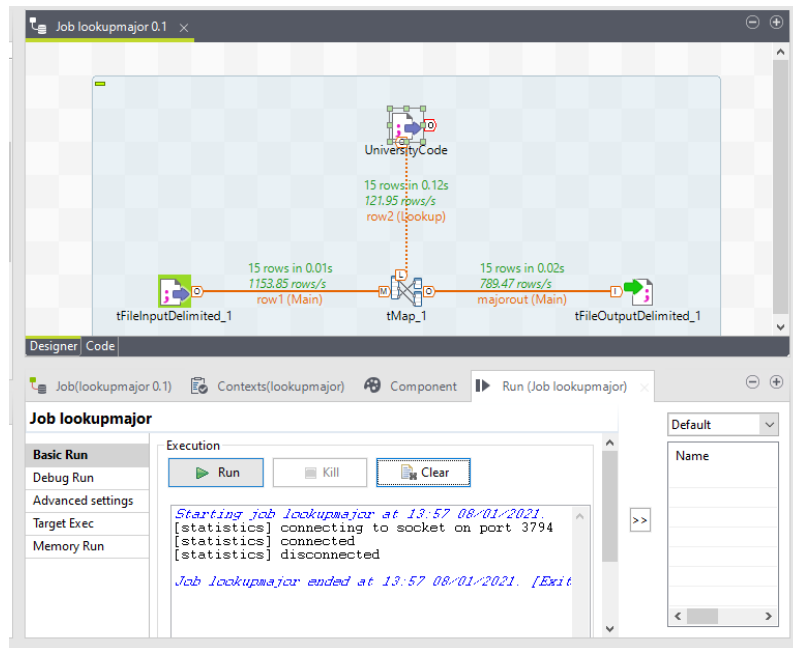


7. Kemudian ubah nama ouput menjadi **universityout.csv** di tFileOutputDelimited



Kemudian untuk melihat hasil penggabungan antara University dengan name_university lakukan perintah **Run**.

Hasil run berhasil tanpa adanya error:



- Hasil dari file universityout.csv kode universitas dengan nama universitas berhasil tergabung:

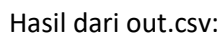
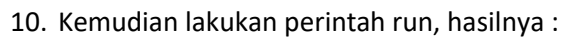
The screenshot shows a Microsoft Excel spreadsheet titled 'universityout - Microsoft Excel (Product Activation Failed)'. The spreadsheet contains a table with 16 rows of data. The first row is the header, and the subsequent rows contain student information.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	First,Last,Gender,Degree,Major,University												
2	Dina,Amelia,Female,S1,PENDIDIKAN DOKTER,UI	-	Universitas Indonesia										
3	Andin,Permata,Female,S1,AKUNTANSI,UB	-	Universitas Brawijaya										
4	Bianca,Afriza,Female,S1,HUBUNGAN INTERNASIONAL,UGM	-	Universitas Gajah Mada										
5	Cindy,Syaputri,Female,S1,SASTRA INGGRIS,IPB	-	Institut Pertanian Bogor										
6	Chandra,Kirana,Female,S1,PSIKOLOGI,UNY	-	Universitas Negeri Yogyakarta										
7	Devina,Anggraini,Female,S1,ILMU KOMUNIKASI,UPI	-	Universitas Pendidikan Indonesia Bandung										
8	Daniar,Rahmawati,Female,S1,MANAJEMEN,UNDIP	-	Universitas Diponegoro										
9	Erhas,Syandika,Male,S1,ILMU KOMPUTER,UP	-	Universitas Pancasila										
10	Farhan,Febriansyah,Male,S1,KRIMINOLOGI,UNES	-	Universitas Negeri Semarang										
11	Gandha,Raharja,Male,S1,SISTEM INFORMASI,UIN	-	Universitas Islam Indonesia										
12	Handika,Putra,Male,S1,MATEMATIKA,UNPAD	-	Universitas Padjajaran										
13	Intan,Dwirizki,Male,S1,STATISTIKA,UNJ	-	Universitas Negeri Jakarta										
14	Johannes,Hutapea,Male,S1,ILMU HUKUM,UKI	-	Universitas Kristen Indonesia										
15	Kevin,Sanjaya,Male,S1,MANAJEMEN,UPH	-	Universitas Pelita Harapan										
16	Lukman,Hardian,Male,S1,BIOLOGI,UNHAS	-	Universitas Hasanuddin										

- Kemudian lakukan penggabungan 2 kolom **First di row1** dan Last di tMap -> drag **Last di row1** -> ke **row1.First** di majorout .

Dengan format:

`row1.First + " " + row1.Last` ... First

[illegible]

[illegible][illegible][illegible]