

# District Data Coordinator Toolbox: Automating Data Acquisition Using Database Connections in SPSS

Jason Schoeneberger, Ph.D.  
Senior Researcher & Task Lead



**Mid-Atlantic:** Delaware, Maryland,  
New Jersey, Pennsylvania, Washington, D.C.

## Prerequisite

---

For this presentation, we assume you have an established database connection.

If not, please review the tool below, which is available on the REL Mid-Atlantic website:

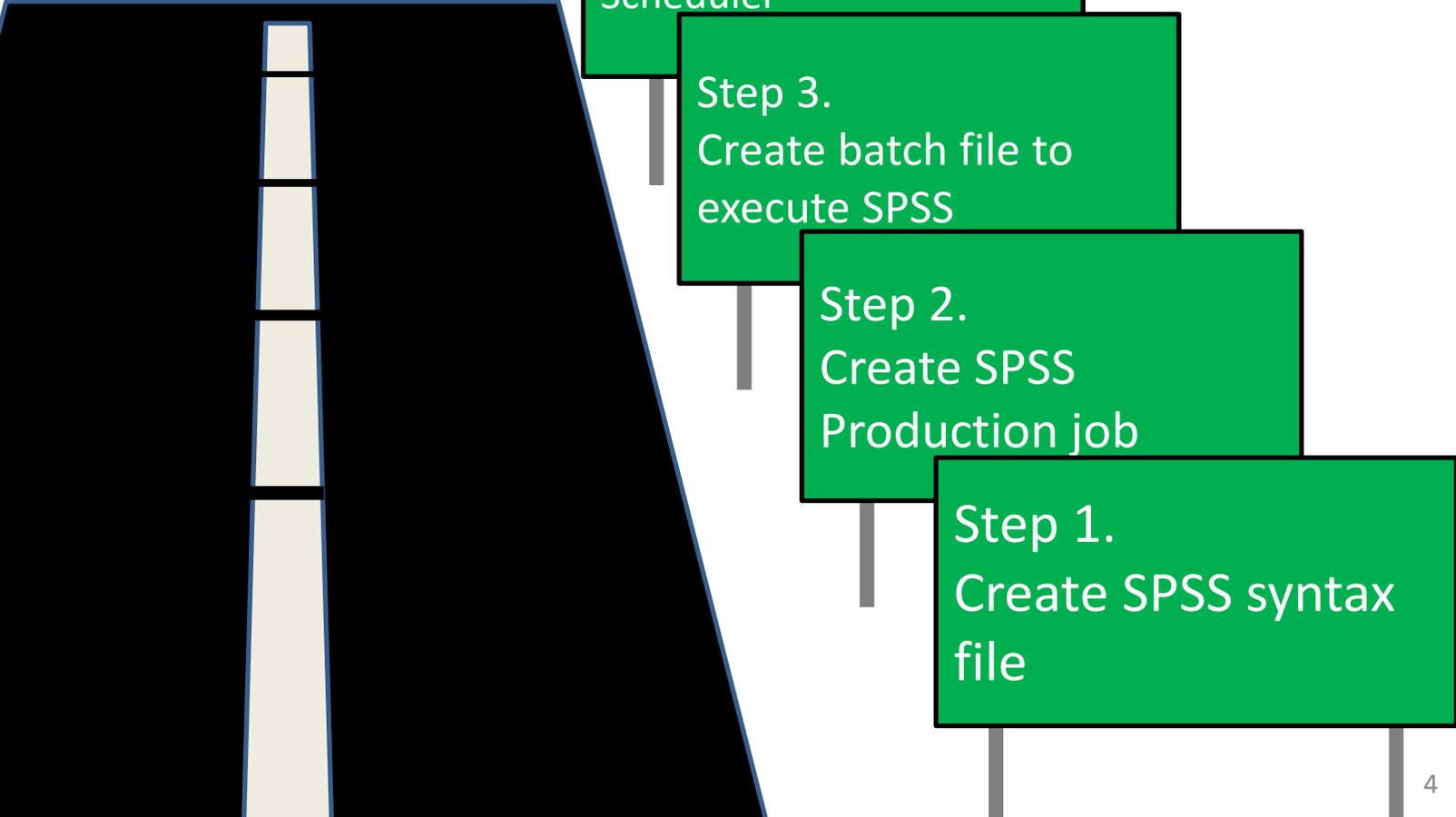
***District Data Coordinator Toolbox:  
Implementing Database Connections  
in Excel***

## Taking the next step...

---

- ■■■ You're making use of a database connection(s) in SPSS, and your increased efficiency resulted in more data requests from stakeholders
- ■■■ Let us suppose some of those data needs are repetitive, or cyclical in nature
  - Maybe someone wants a report updated on a monthly or weekly basis
  - Maybe data in the database is refreshed each evening, and you want the latest available to you each morning
- ■■■ Let's automate that data acquisition process!

# Road map to data connectivity



Step 4.  
Schedule task with Task  
Scheduler

Step 3.  
Create batch file to  
execute SPSS

Step 2.  
Create SPSS  
Production job

Step 1.  
Create SPSS syntax  
file

## Our road map to automation

---

1. We prepare an SPSS syntax file to generate what we need
2. We create an SPSS Production Job to execute our syntax file
3. A batch file is created containing instructions to tell our computer to execute our SPSS Production job
4. Then we schedule a task (running the batch file) using the Task Scheduler

## Traveling the road by example

---

- To follow the steps in our road map to connectivity, let's assume the following example:
  - District leadership is focused on monitoring student mobility
  - As a result, several principals with highly-mobile populations have requested a weekly summary report of enrollment at their middle schools
  - They want to examine enrollment, disaggregated by grade level and student race/ethnicity
  - The data we need to obtain are stored in an Access database

# Report to be automated

- Below is the table we want to automatically refresh each week.

			student_grade			Total
			06	07	08	
student_ethnicity_desc	African American	Count	132	176	188	496
		% within student_ethnicity_desc	26.6%	35.5%	37.9%	100.0%
		% within student_grade	53.4%	34.0%	34.6%	37.9%
	American Indian	Count	1	2	2	5
		% within student_ethnicity_desc	20.0%	40.0%	40.0%	100.0%
		% within student_grade	0.4%	0.4%	0.4%	0.4%
	Asian	Count	7	13	13	33
		% within student_ethnicity_desc	21.2%	39.4%	39.4%	100.0%
		% within student_grade	2.8%	2.5%	2.4%	2.5%
Hispanic	Count	54	74	67	195	
	% within student_ethnicity_desc	27.7%	37.9%	34.4%	100.0%	
	% within student_grade	21.9%	14.3%	12.3%	14.9%	
Multi-Racial	Count	6	13	6	25	
	% within student_ethnicity_desc	24.0%	52.0%	24.0%	100.0%	
	% within student_grade	2.4%	2.5%	1.1%	1.9%	
White	Count	47	240	268	555	
	% within student_ethnicity_desc	8.5%	43.2%	48.3%	100.0%	
	% within student_grade	19.0%	46.3%	49.3%	42.4%	
Total	Count	247	518	544	1309	
	% within student_ethnicity_desc	18.9%	39.6%	41.6%	100.0%	
	% within student_grade	100.0%	100.0%	100.0%	100.0%	

## Using SPSS syntax

---

- To automate the refresh of our table, we need to use SPSS syntax to accomplish our tasks

# Compiling SPSS commands in syntax form

1. Obtain data to work with (here we are querying a database)

2. Analyze or manipulate data in logical order to generate desired results (here, a crosstab table)

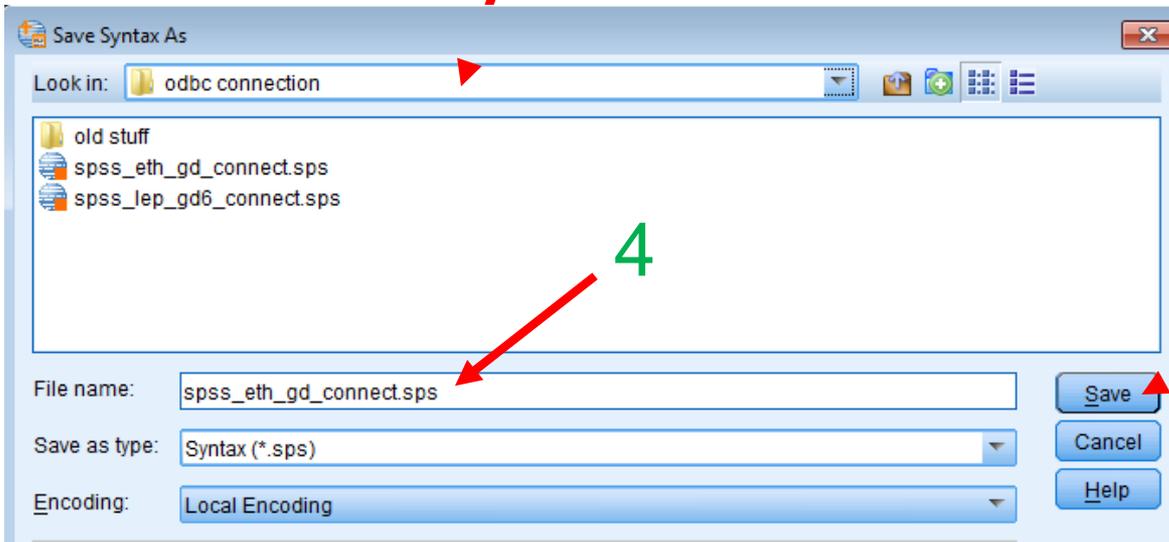
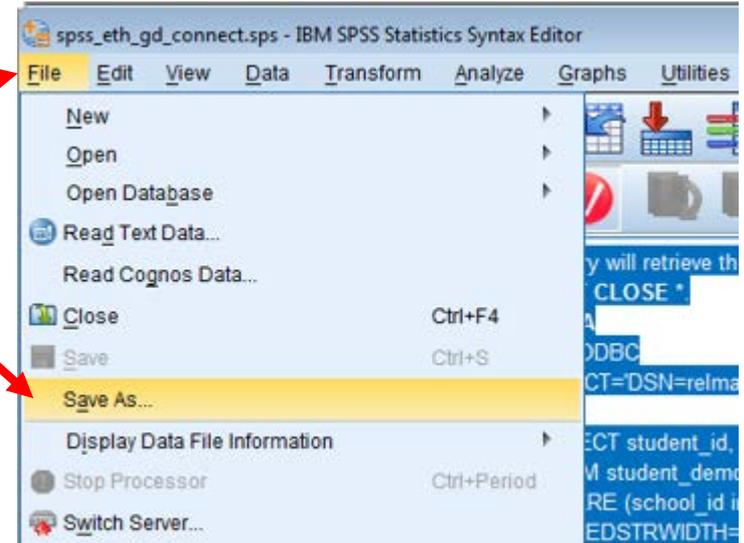
3. Save data file

4. Save output

```
1  **this query will retrieve the data necessary to create our summary table.
2  DATASET CLOSE *.
3  GET DATA
4  /TYPE=ODBC
5  /CONNECT='DSN=relma_access_connect;DBQ=C:\Users\Jason\Desktop\connection_test\connect_training.accdb;Driver=
6  /SQL=
7  "SELECT student_id, school_id, student_grade, student_ethnicity_desc "+
8  "FROM student_demographics "+
9  "WHERE (school_id in(301,305,313))"
10 /ASSUMEDSTRWIDTH=255.
11 CACHE.
12 EXECUTE.
13
14 **create crosstabulation table, with percentages by row and column.
15 CROSSTABS
16 /TABLES=student_ethnicity_desc BY student_grade
17 /FORMAT=AVALUE TABLES
18 /CELLS=COUNT ROW COLUMN
19 /COUNT ROUND CELL.
20
21 **save data file.
22 SAVE OUTFILE='C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc connection\gd_eth_data.sav'
23 /COMPRESSED.
24
25 **save output.
26 OUTPUT SAVE NAME=Document1 OUTFILE='C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc '+
27 'connection\gd_eth_output.spv' LOCK=NO.
```

# Save syntax file

1. Click File in the syntax window
2. Click Save As...
3. Navigate to folder location
4. Name syntax file
5. Click Save



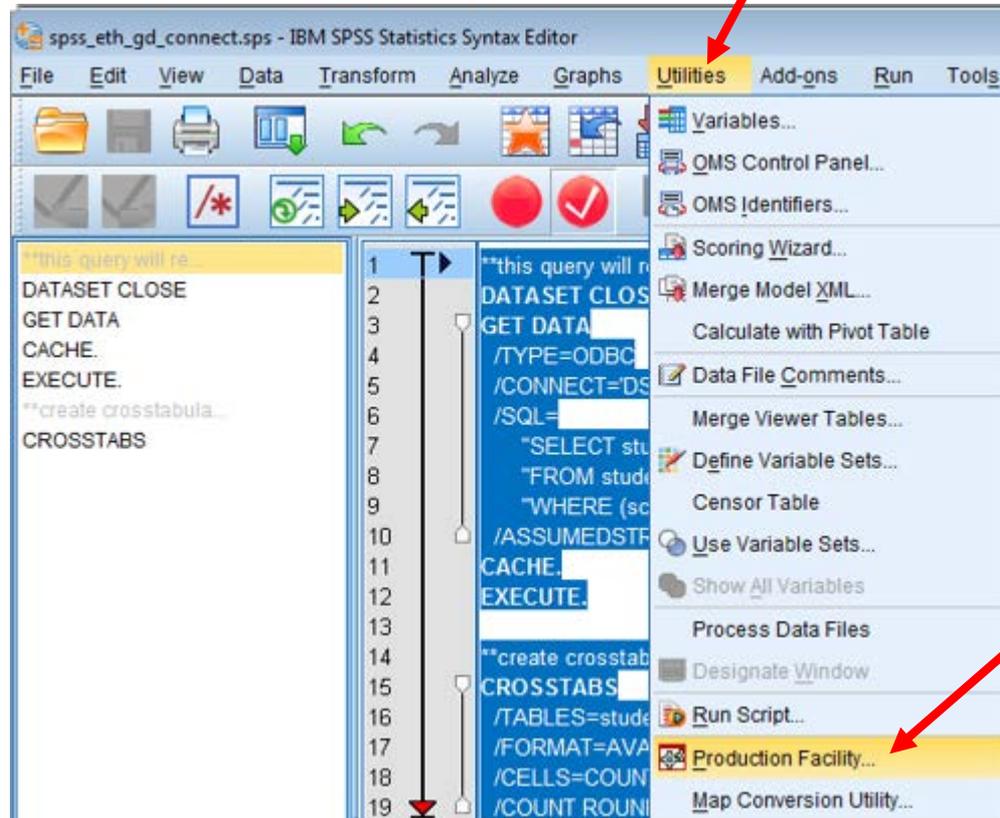
# Using SPSS production facility

---

- SPSS provides a production facility for the purposes of automating work. Let's use that to set up a structure for executing syntax and saving output

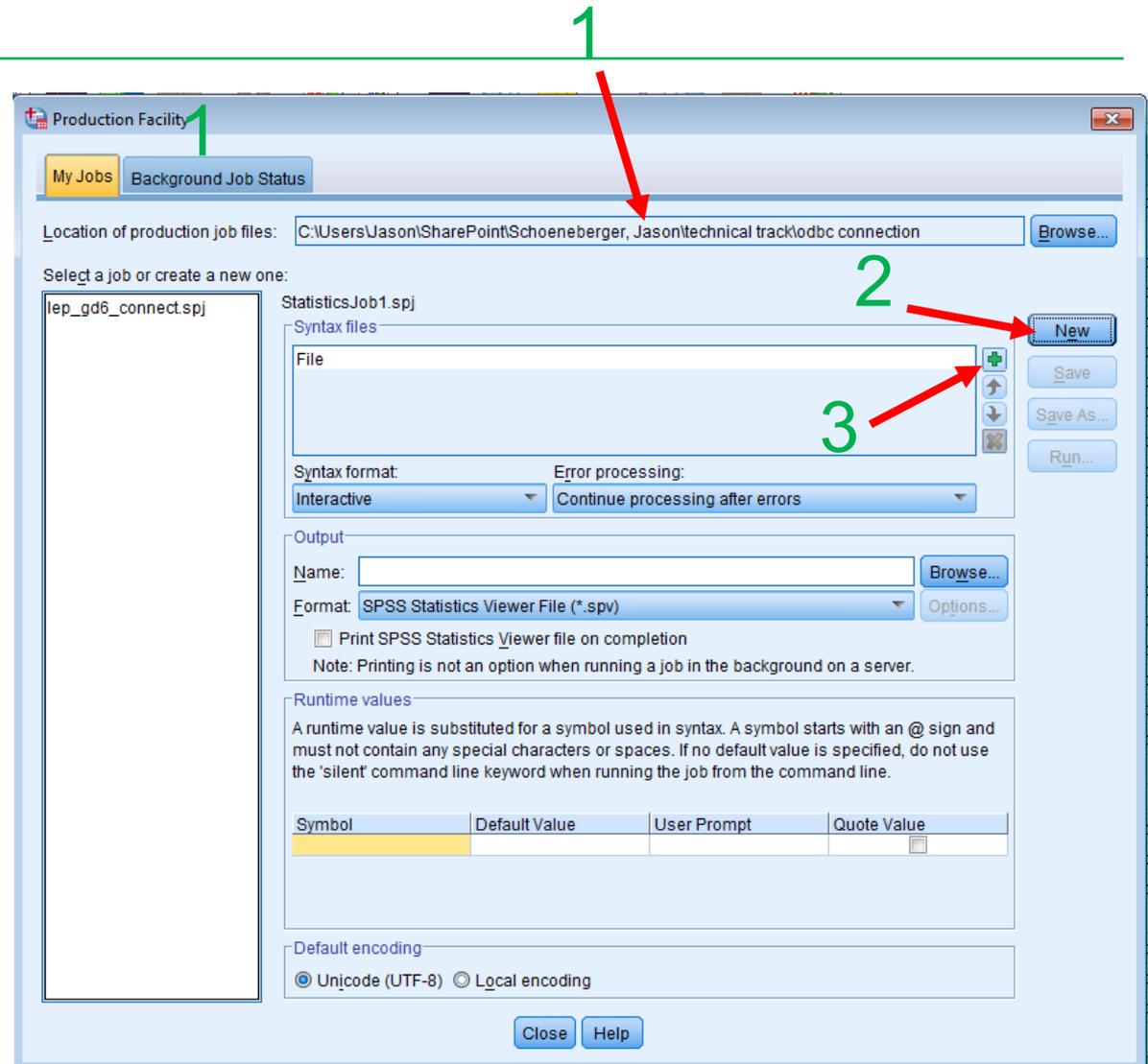
# Create SPSS production job

1. Click Utilities
2. Click Production Facility...



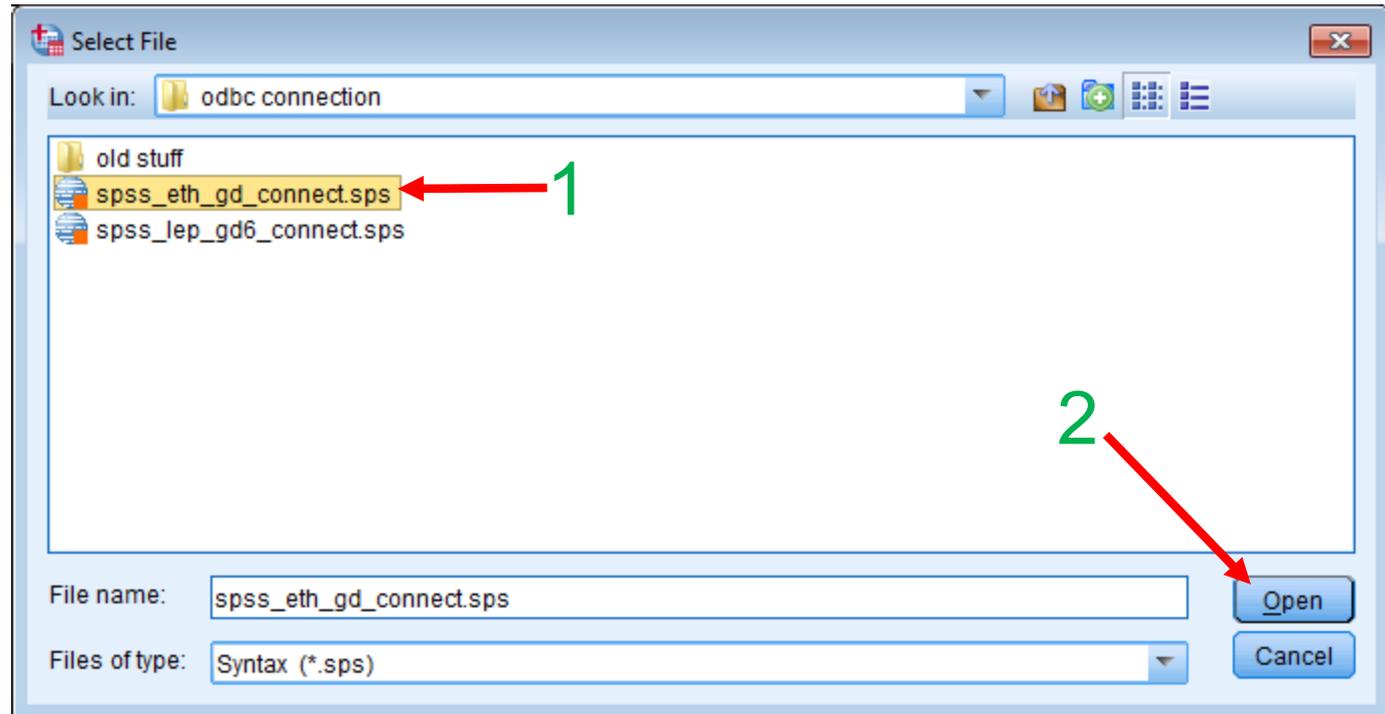
# Specify location for production file

1. Use navigation bar to specify location for production job file
2. Click New to add a syntax file to the job
3. Click green plus button



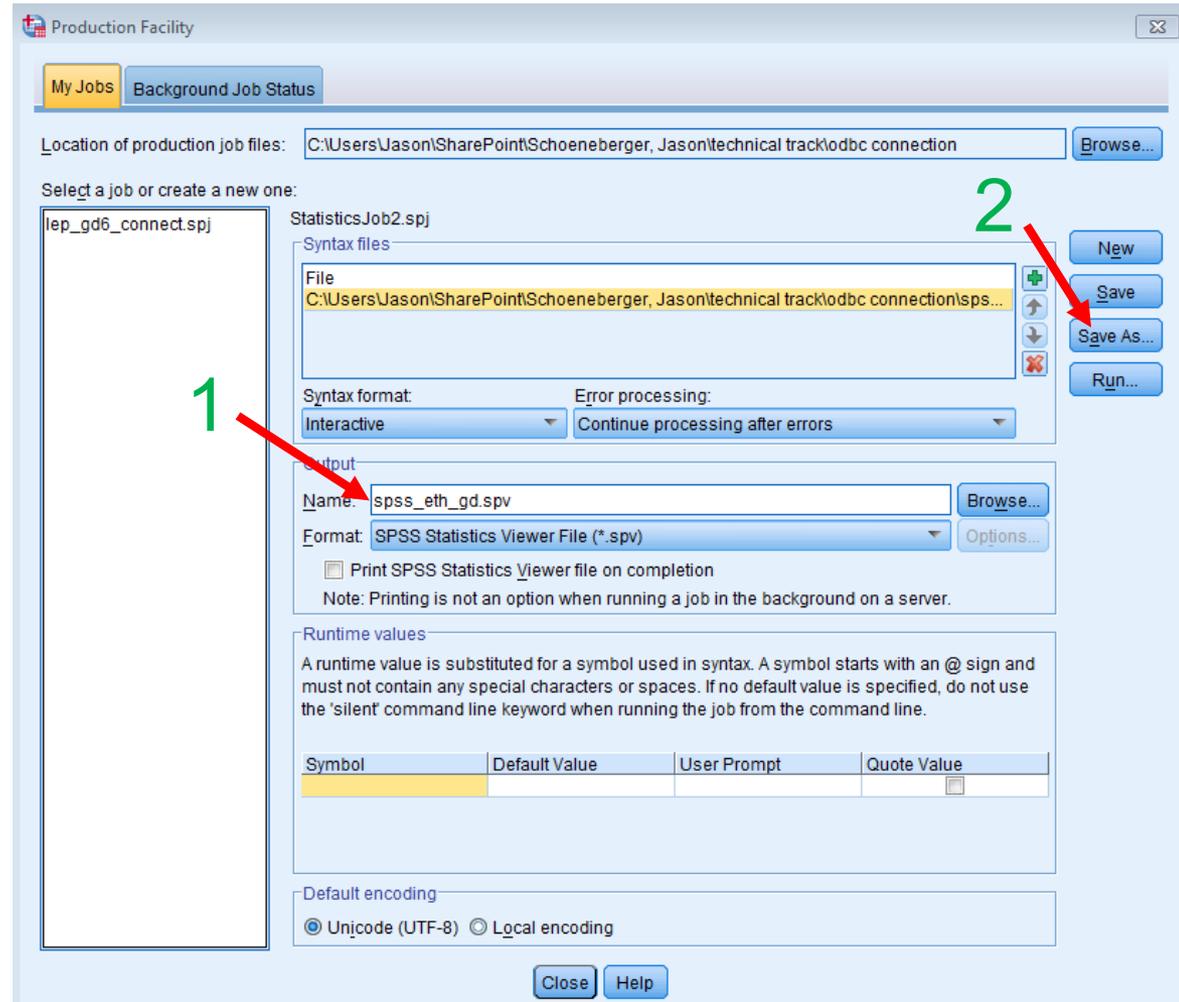
# Select syntax file to execute

1. Select syntax file to execute
2. Click Open



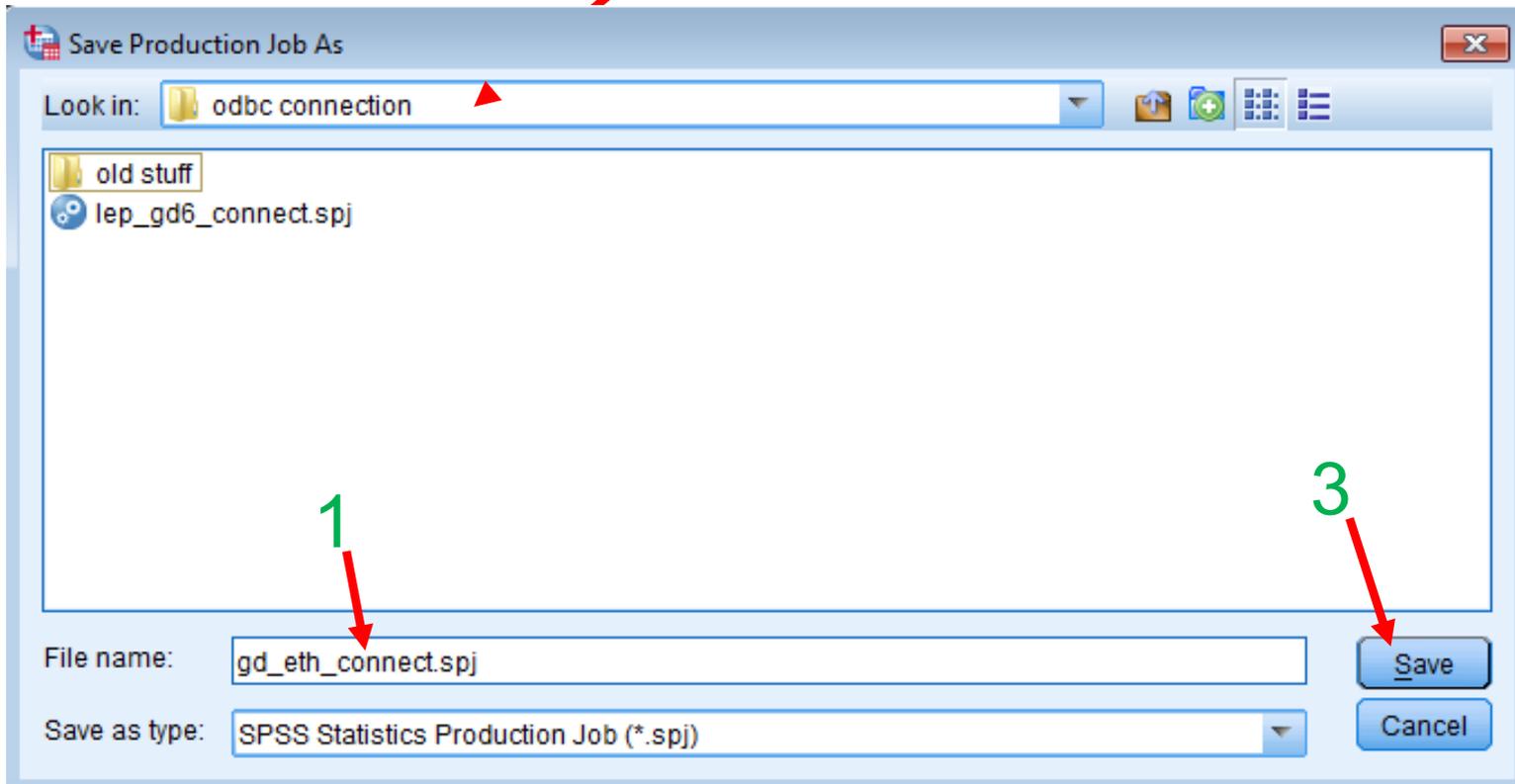
# Name SPSS output file

1. Name the SPSS output file generated by the execution of the syntax file
2. Click Save As



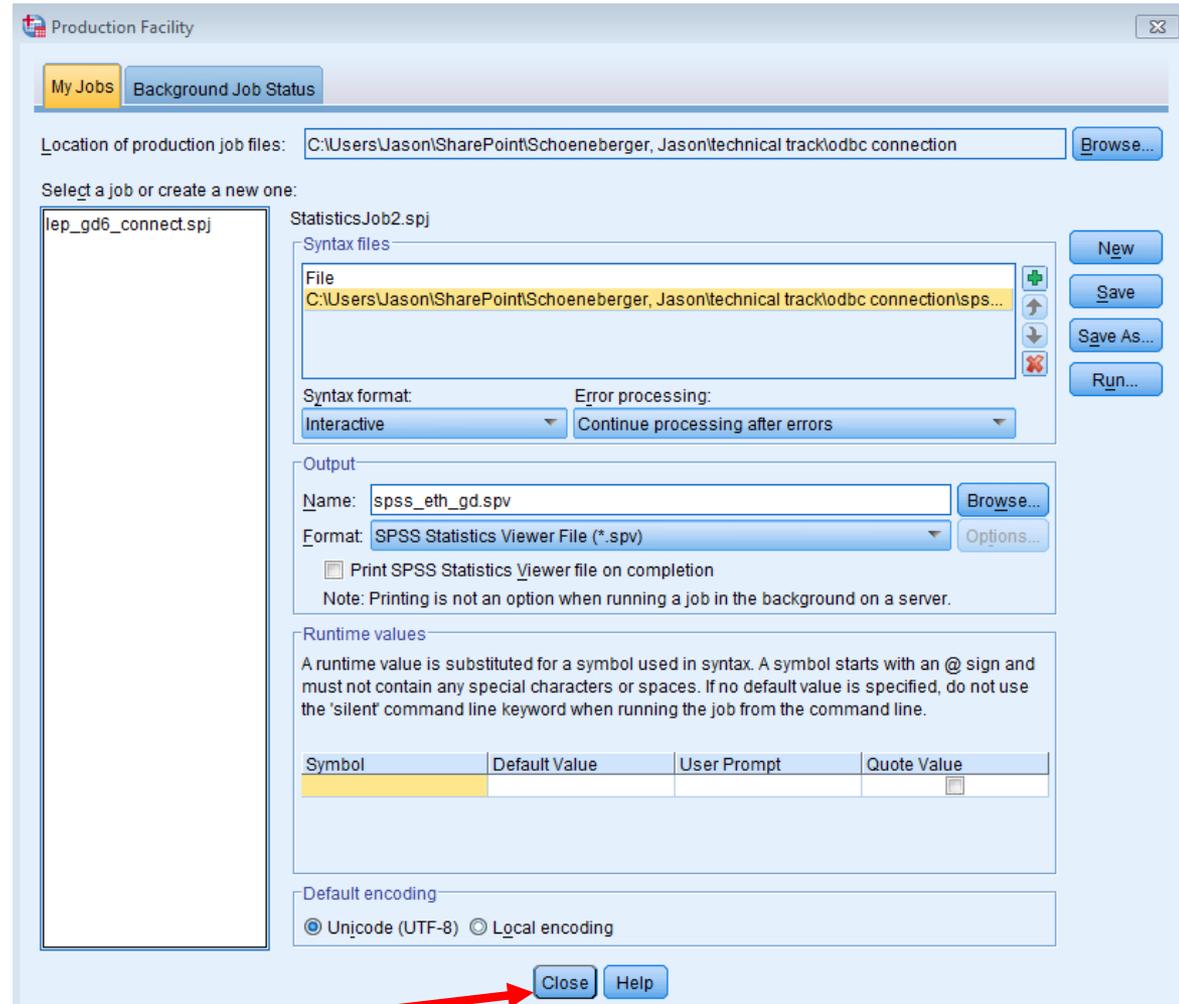
# Name SPSS production job

1. Name the SPSS production job
2. Specify location
3. Click Save



# Close production facility

## 1. Close production facility



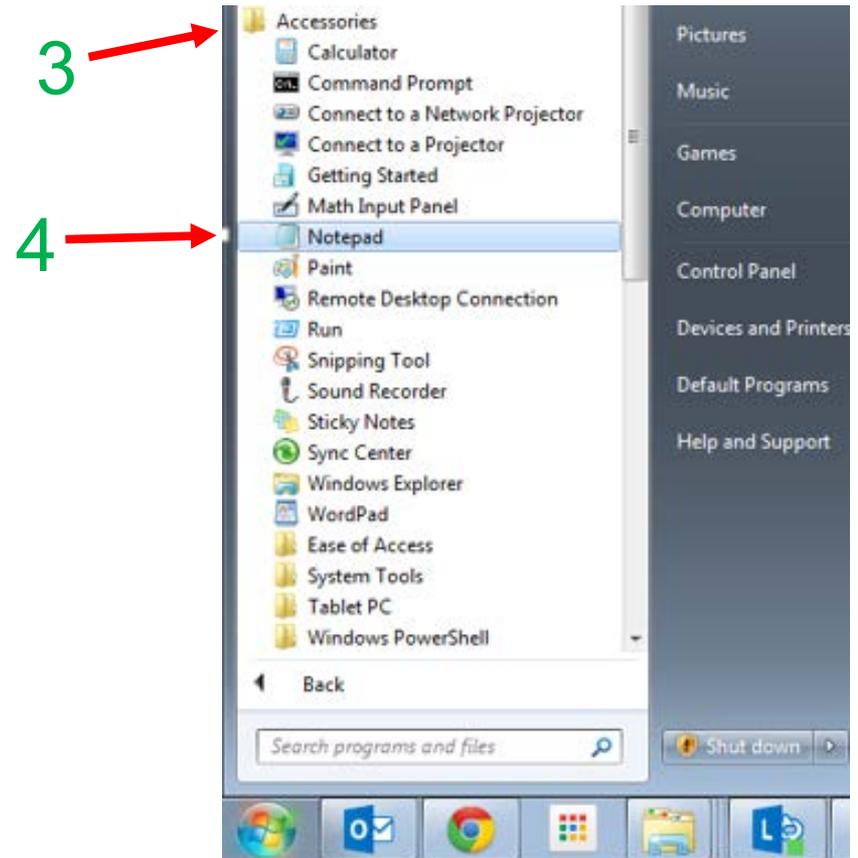
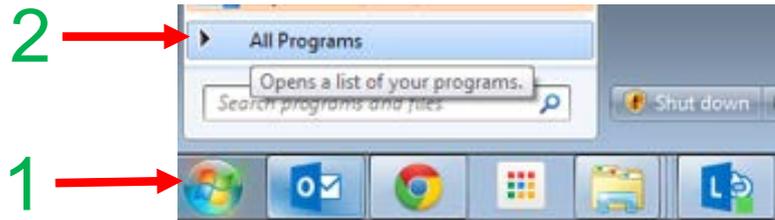
## Creating a batch file

---

- Using Notepad or another text editor, we can create a batch (.bat extension) file containing instructions informing our computer to execute our SPSS production job

# Opening Notepad

1. Click on the Windows icon
2. Click All Programs
3. Click on Accessories
4. Click on Notepad



# Batch file code for copy-pasting

---

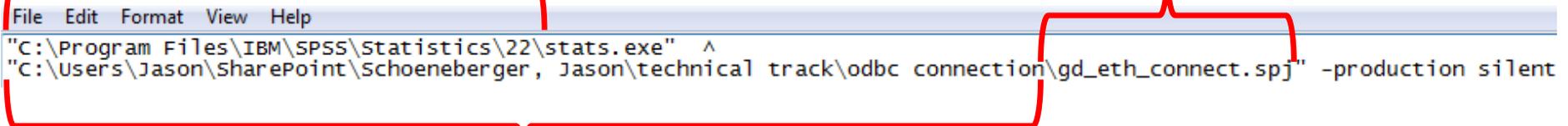
- Below is the code to be copy-pasted into the batch file

```
"C:\Program Files\IBM\SPSS\Statistics\22\stats.exe" ^
```

```
"C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc  
connection\gd_eth_connect.spj" -production silent
```

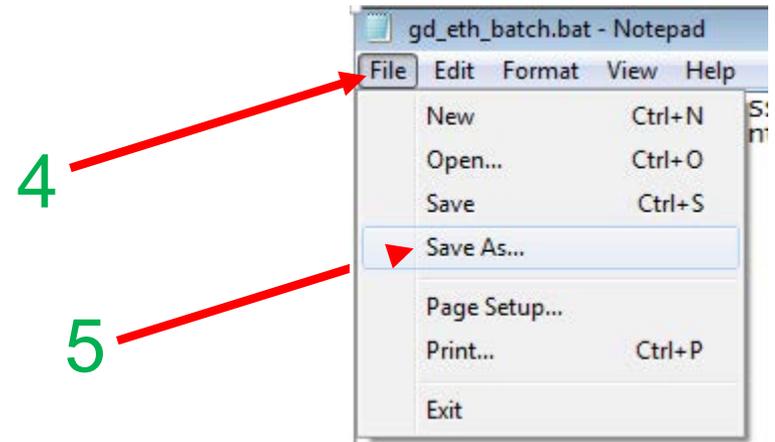
## Creating the batch file

1. Specify the specific folder path for the SPSS executable file
2. Specify the path for the saved SPSS production job
3. Specify the saved SPSS production job name



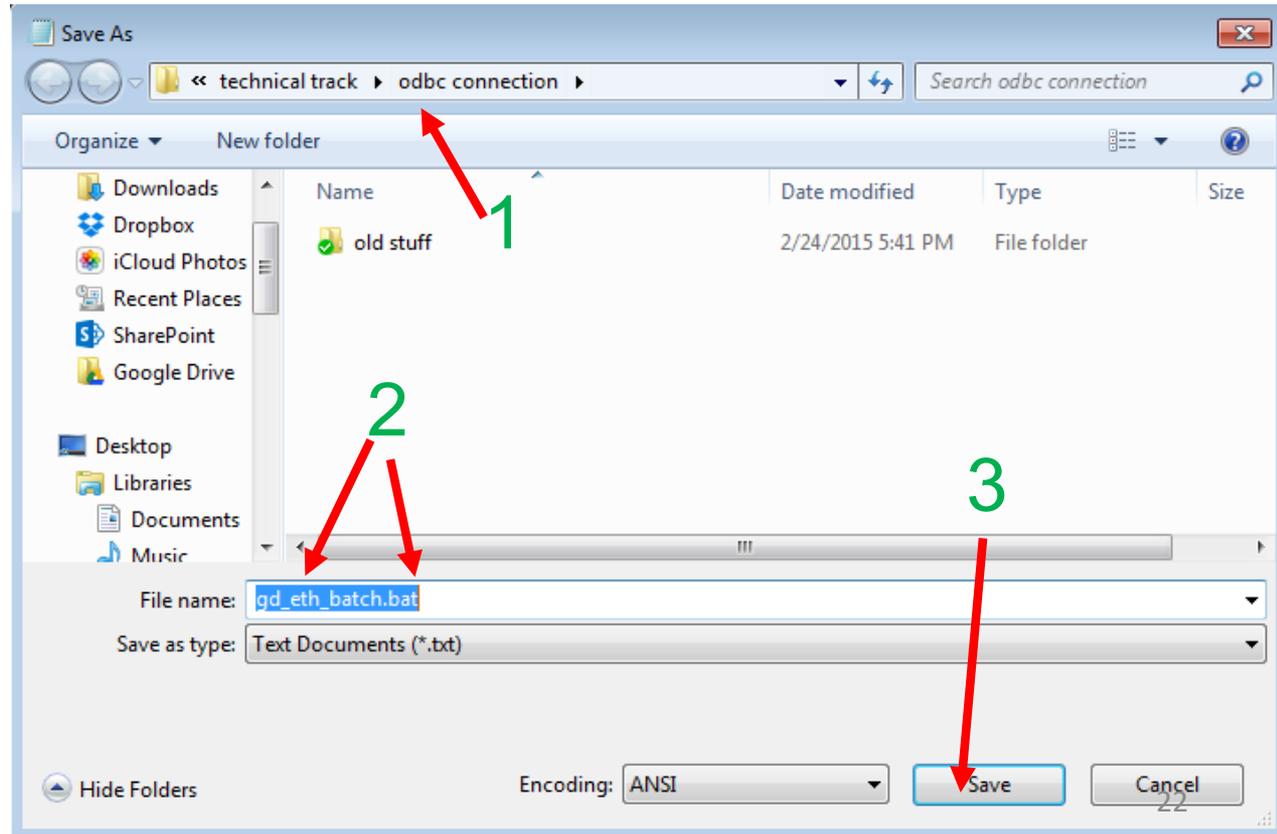
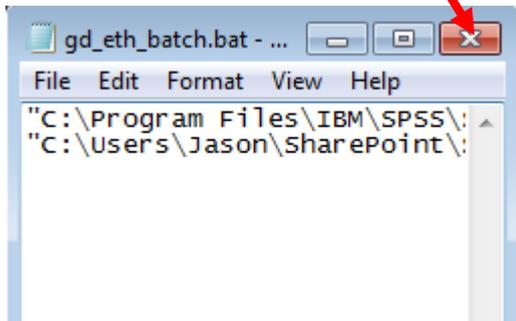
```
File Edit Format View Help  
"C:\Program Files\IBM\SPSS\Statistics\22\stats.exe" ^  
"C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc connection\gd_eth_connect.spj" -production silent
```

4. Click File
5. Click Save As



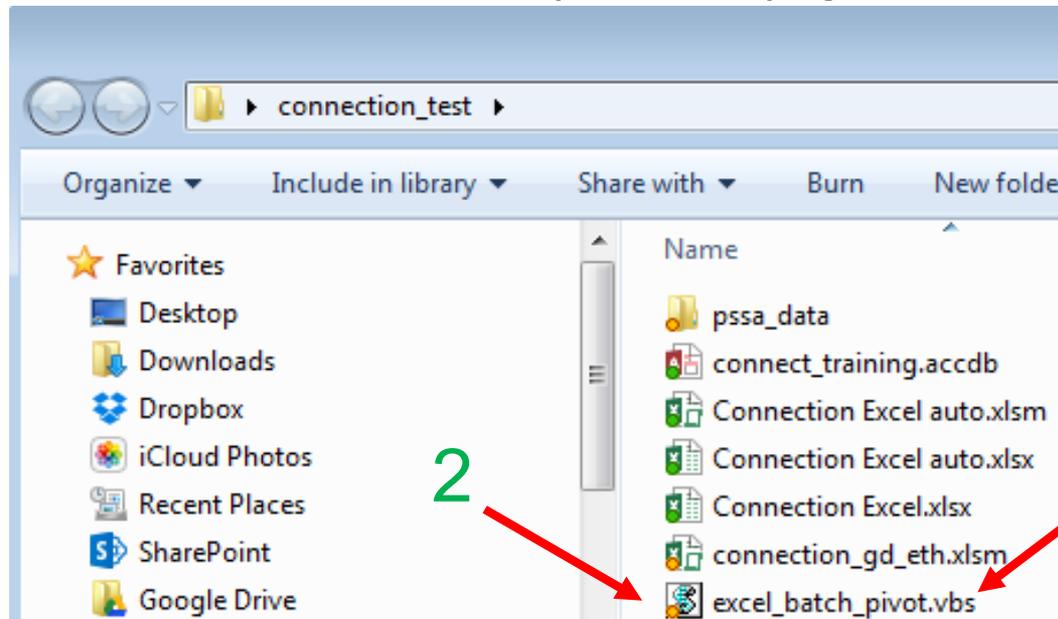
# Saving the batch file

1. Navigate to your chosen folder
2. Name your batch file (gd\_eth\_batch), and enter '.bat' as the file extension
3. Click Save
4. Close Notepad



## Saving the batch file

1. Verify batch file was saved
2. You can check that it works by double-clicking on the batch file itself
3. You should see SPSS open briefly, then immediately close
4. If there is an error in the code, you may get an error notice

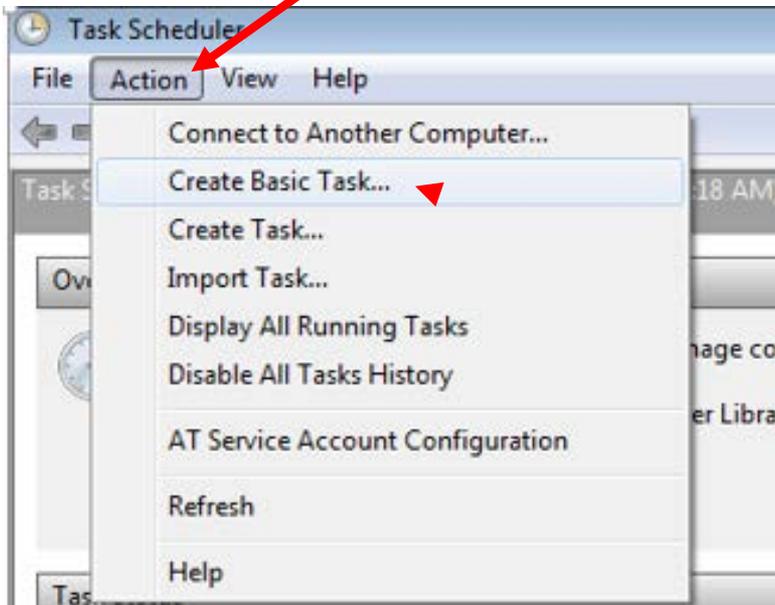
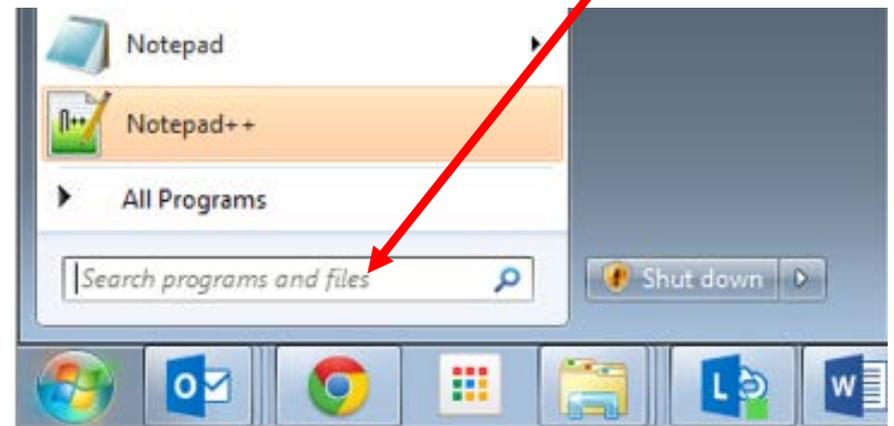


## Scheduling a task

- Now we can use the Task Scheduler to run our batch file, which executes the SPSS syntax file and refreshes our output crosstab table

# Opening Task Scheduler

1. Enter 'Task Scheduler' in search box
2. Click on Action in the Task Scheduler window
3. Select Create Basic Task



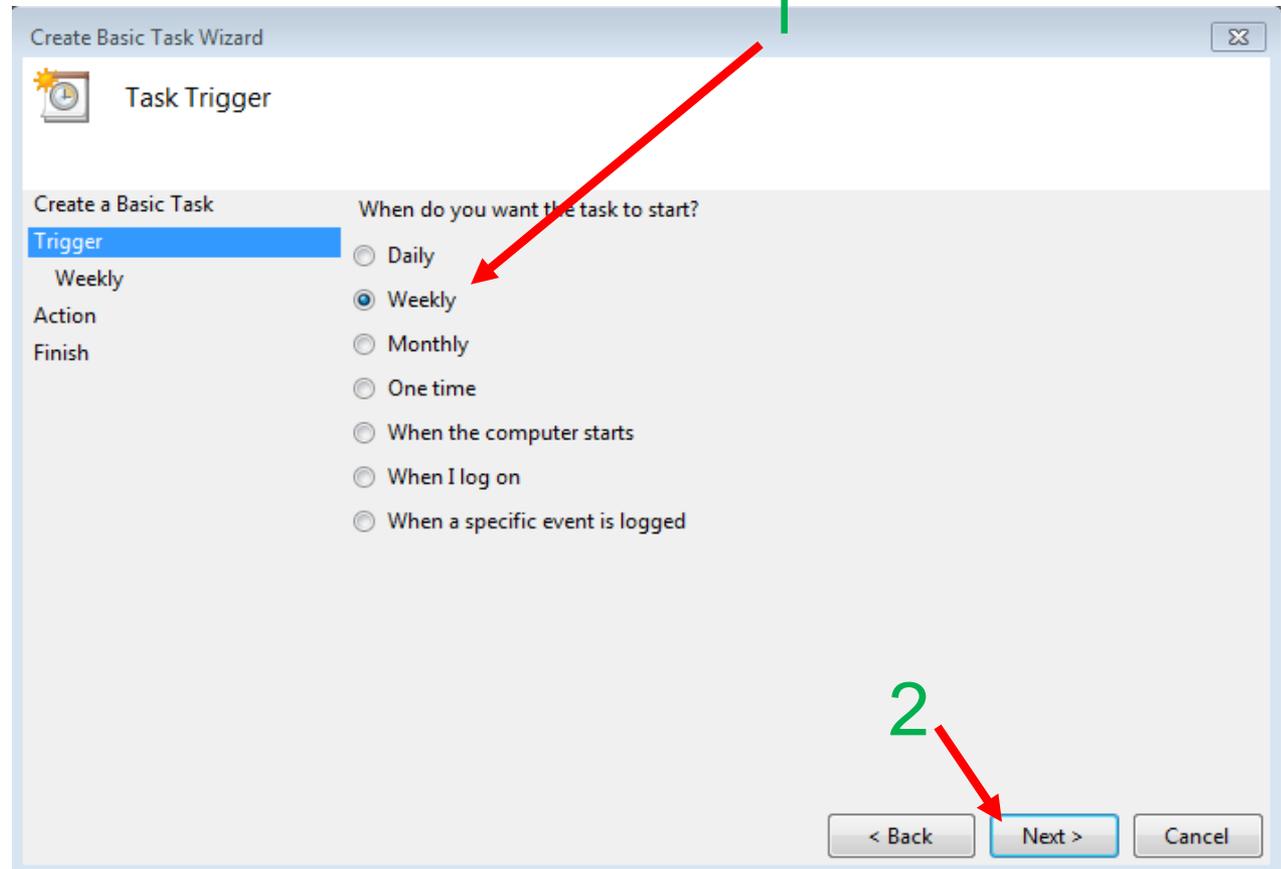
# Naming the scheduled task

1. Enter a name for the scheduled task  
(e.g gd\_eth\_spssprod\_refresh)
2. Click Next

The screenshot shows the 'Create Basic Task Wizard' dialog box. The title bar reads 'Create Basic Task Wizard'. Below the title bar, there is a clock icon and the text 'Create a Basic Task'. The main area is divided into a left sidebar and a right main area. The sidebar has four items: 'Create a Basic Task' (highlighted in blue), 'Trigger', 'Action', and 'Finish'. The main area contains the text: 'Use this wizard to quickly schedule a common task. For more advanced options or settings such as multiple task actions or triggers, use the Create Task command in the Actions pane.' Below this text, there is a 'Name:' label followed by a text input field containing 'gd\_eth\_spssprod\_refresh'. Below the input field is a 'Description:' label followed by a large empty text area. At the bottom right, there are three buttons: '< Back', 'Next >', and 'Cancel'. A green '1' is positioned above the 'Name' input field. A red arrow points from a green '2' to the 'Next >' button.

# Set frequency of scheduled task

1. Select how often you would like the refresh to occur (e.g. a weekly refresh for our principals in this example)
2. Click Next



# Set time-of-day and recurrence of scheduled task

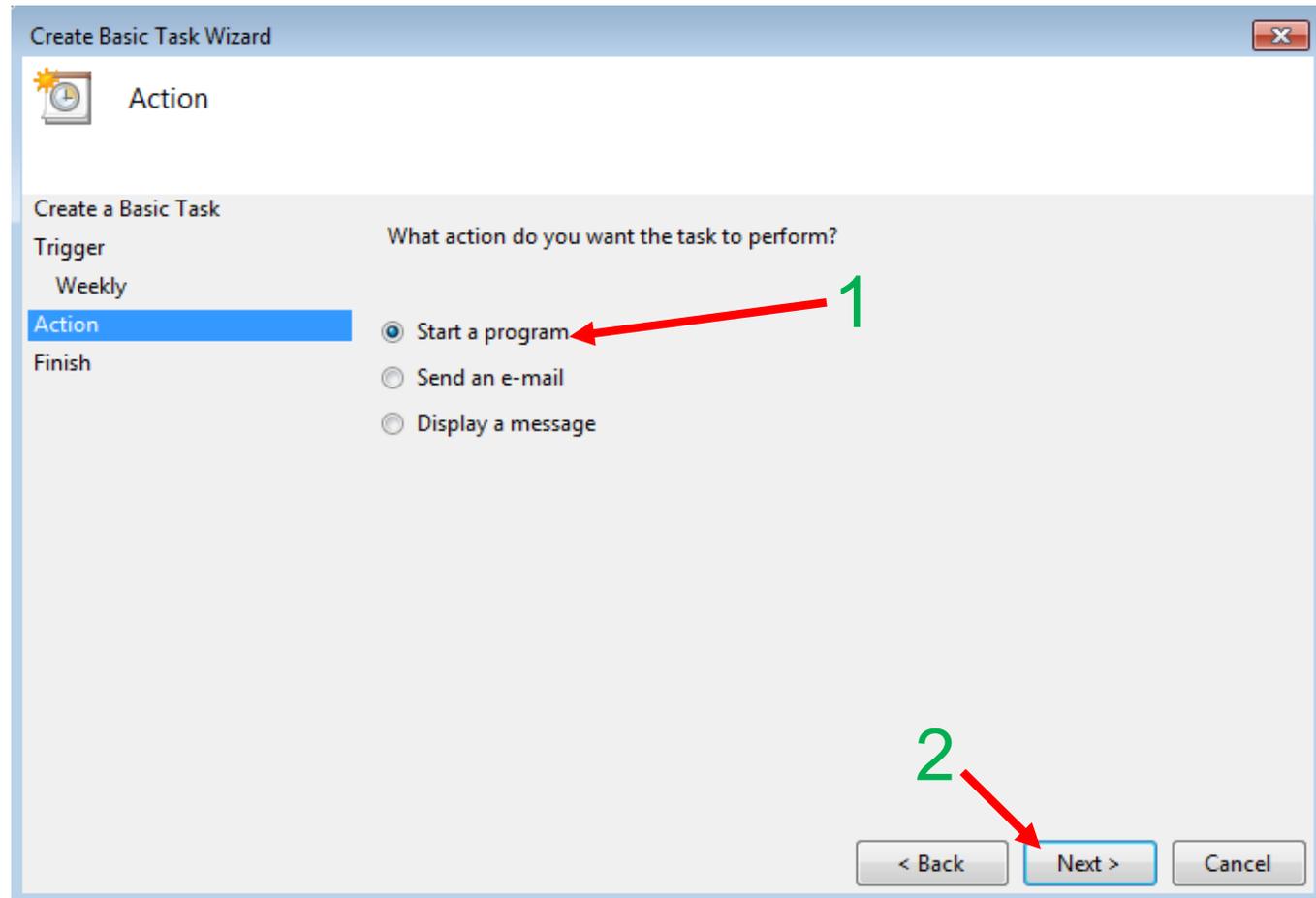
1. Specify the date to begin and time of day you would like to use
2. Specify the day of the week (e.g Friday) you want the refresh to occur
3. Click Next

The screenshot shows the 'Create Basic Task Wizard' dialog box with the 'Weekly' trigger selected. The configuration is as follows:

- Start:** 2/27/2015 (indicated by arrow 1)
- Time:** 7:05:35 AM (indicated by arrow 1)
- Recur every:** 1 weeks on:
- Days:**  Sunday,  Monday,  Tuesday,  Wednesday,  Friday (indicated by arrow 2),  Saturday
- Buttons:** < Back, Next > (indicated by arrow 3), Cancel

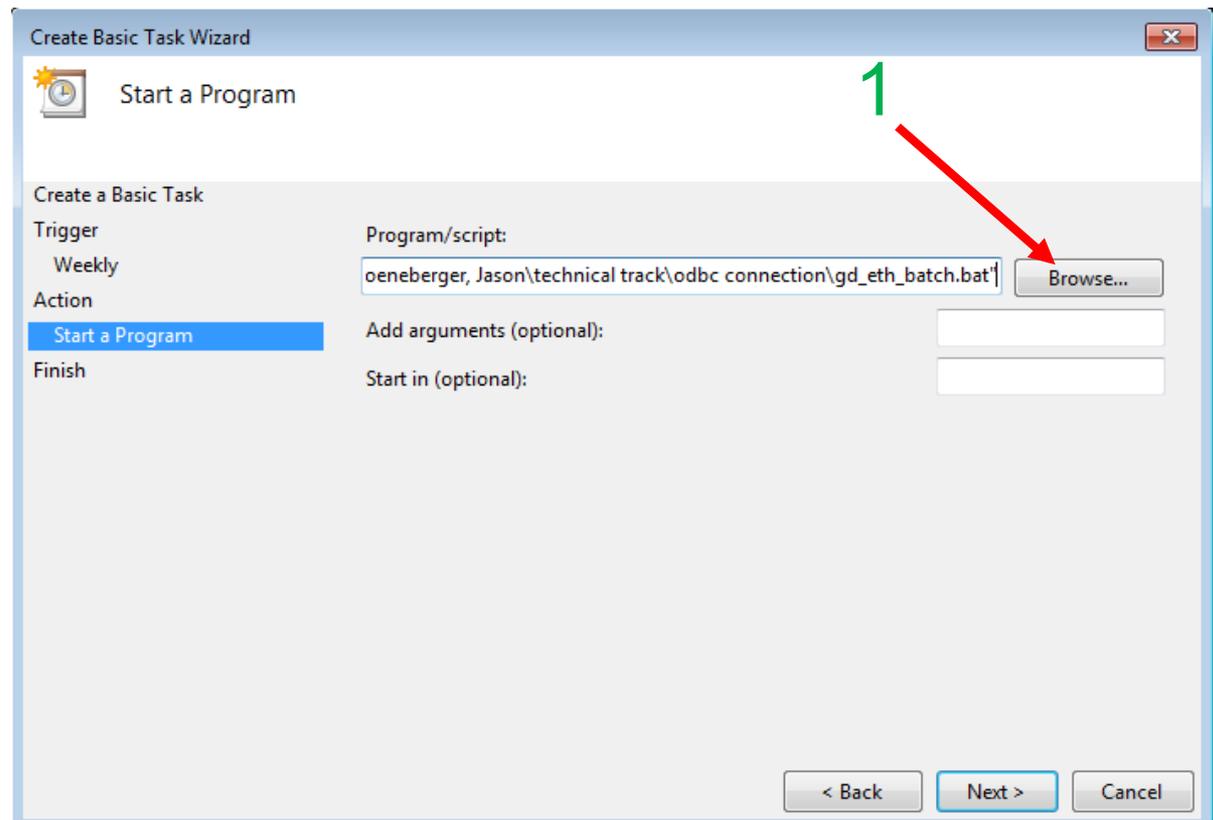
# Specify the type of action to be scheduled

1. Select 'Start a program'
2. Click Next



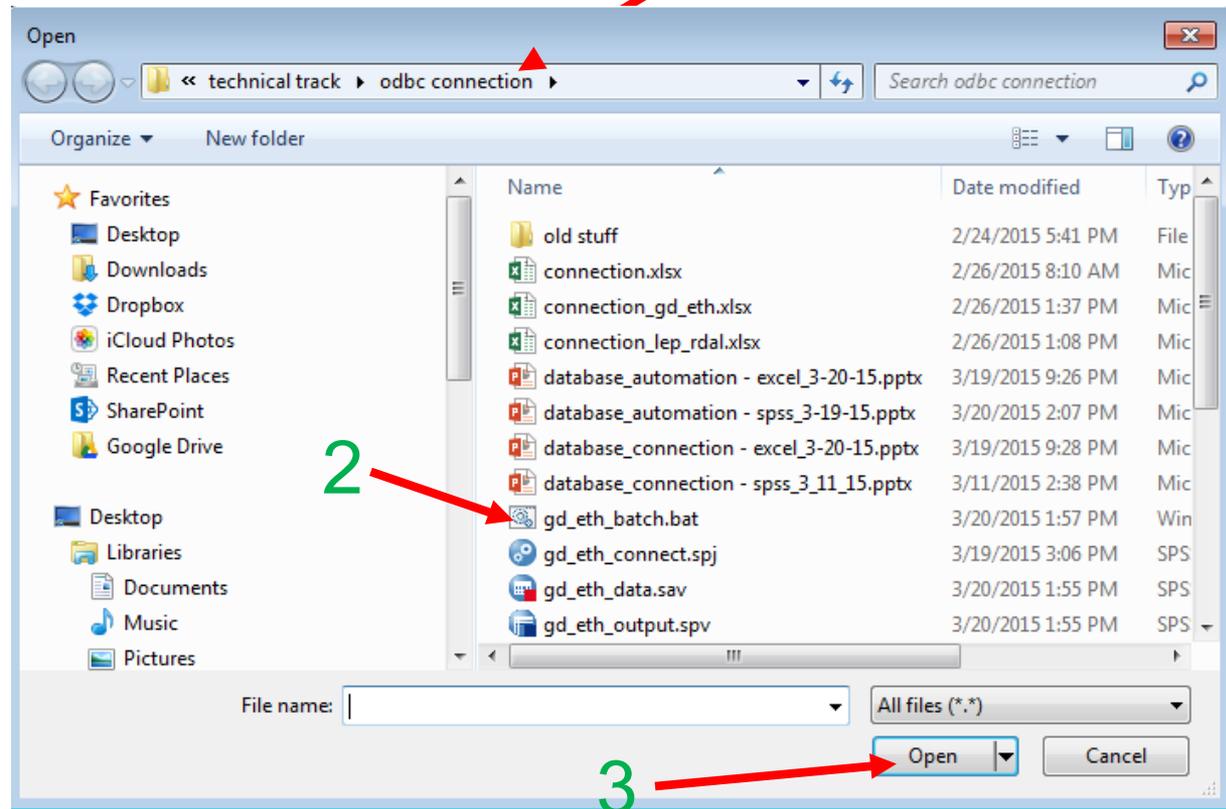
# Specify file scheduled task should execute

1. Click 'Browse' to navigate to folder location where '.bat' file is stored



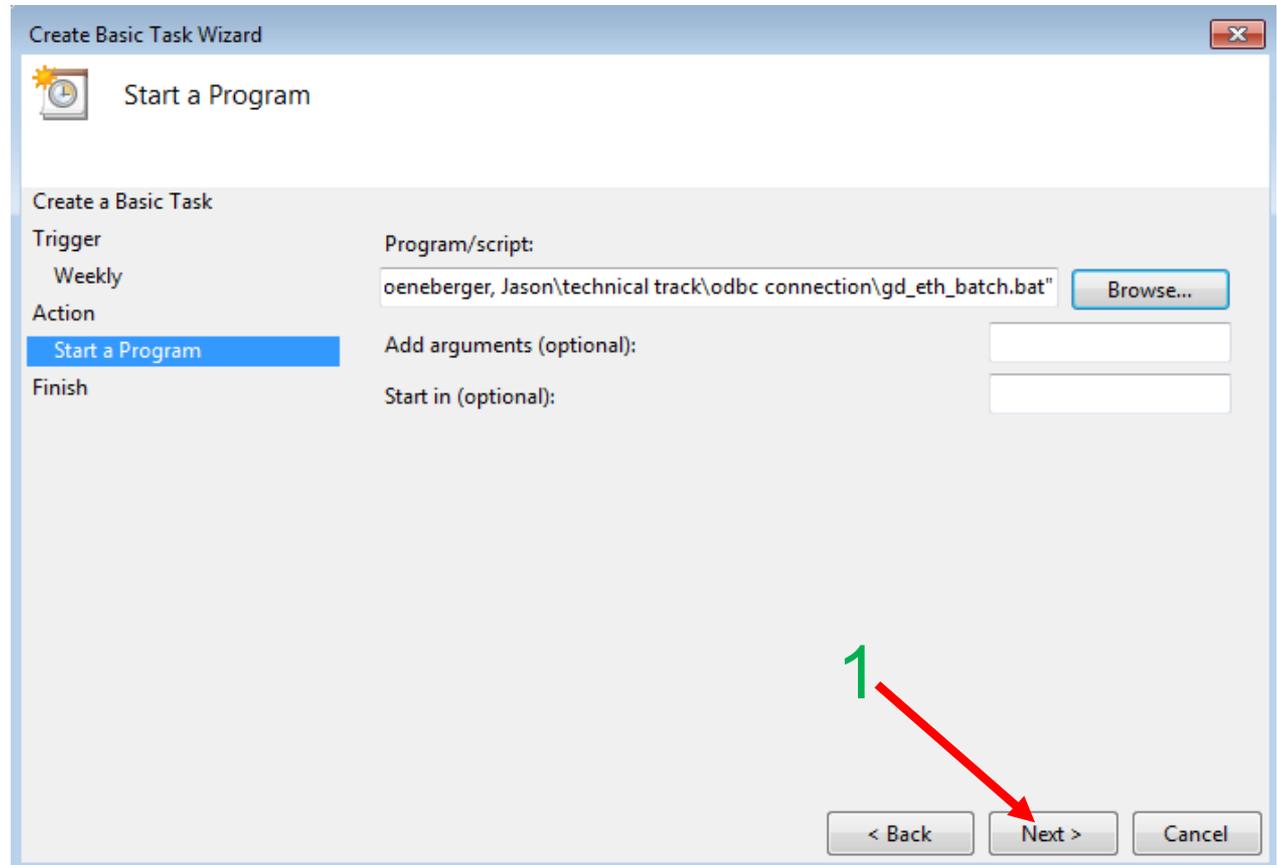
# Specify batch file to execute

1. Navigate to folder location where '.bat' file is stored
2. Select the '.bat' file of interest
3. Click Open



# Specify file scheduled task should execute

## 1. Click Next



# Review & finish scheduling task

## 1. Review properties of scheduled task:

- a) Name of scheduled task
- b) Trigger time
- c) Action (batch file to execute)

## 2. Click Finish

The screenshot shows the 'Create Basic Task Wizard' window with the 'Summary' tab selected. The 'Name' field is 'gd\_eth\_spssprod\_refresh', the 'Trigger' is 'Weekly; At 2:20 PM every Friday of every week, starting 3/20/2015', and the 'Action' is 't\Schoeneberger, Jason\technical track\odbc connection\gd\_eth\_batch.bat'. The 'Finish' button is highlighted. Red arrows point from green labels '1a', '1b', '1c', and '2' to the respective fields and buttons.

# Monitoring/editing a scheduled task

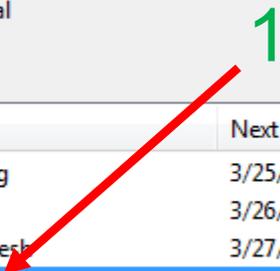
1. Double-click on the scheduled task under the Active Tasks pane inside Task Scheduler:

Active Tasks

Active tasks are tasks that are currently enabled and have not expired.

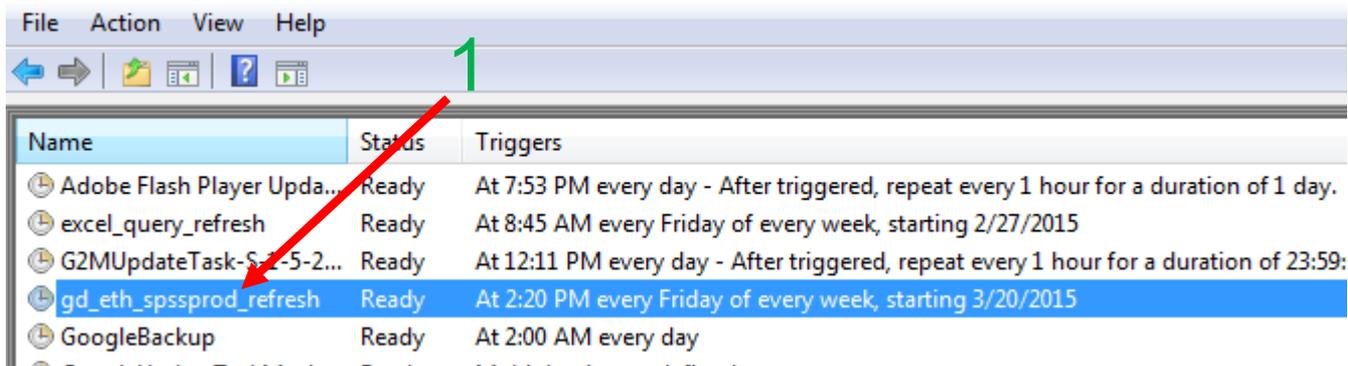
Summary: 61 total

Task Name	Next Run Time	Triggers	Location
ScheduledDefrag	3/25/2015 2:02:05 AM	At 1:00 AM every Wedne...	\Microsoft\Windows\De...
KernelCeipTask	3/26/2015 3:30:00 AM	At 3:30 AM every Thursd...	\Microsoft\Windows\C...
excel_query_refresh	3/27/2015 8:45:00 AM	At 8:45 AM every Friday ...	\
<b>gd_eth_spssprod_refresh</b>	3/27/2015 2:20:29 PM	At 2:20 PM every Friday ...	\
AnalyzeSystem	3/31/2015 1:12:33 PM	At 6:00 AM every 14 days	\Microsoft\Windows\Po...
{78C321B8-66D3-47EF-8D22-404FF...}		When the task is created...	\
{A72B327E-8A2F-49FA-BE35-42718...}		When the task is created...	\

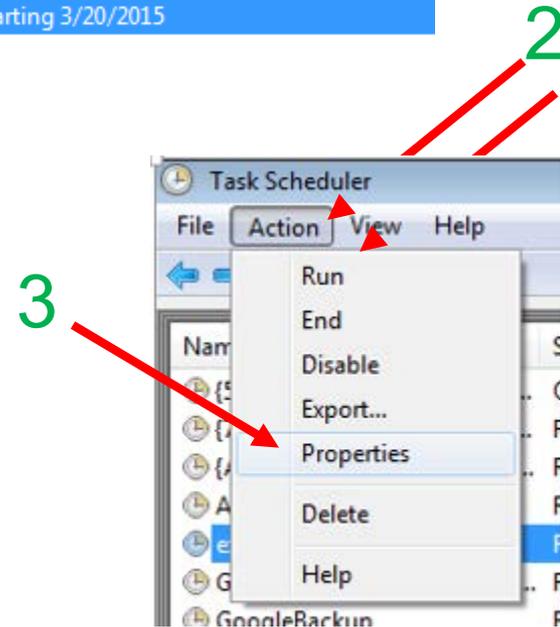


# Monitoring/editing a scheduled task

1. Click on the scheduled task under the Action Pane inside Task Scheduler:

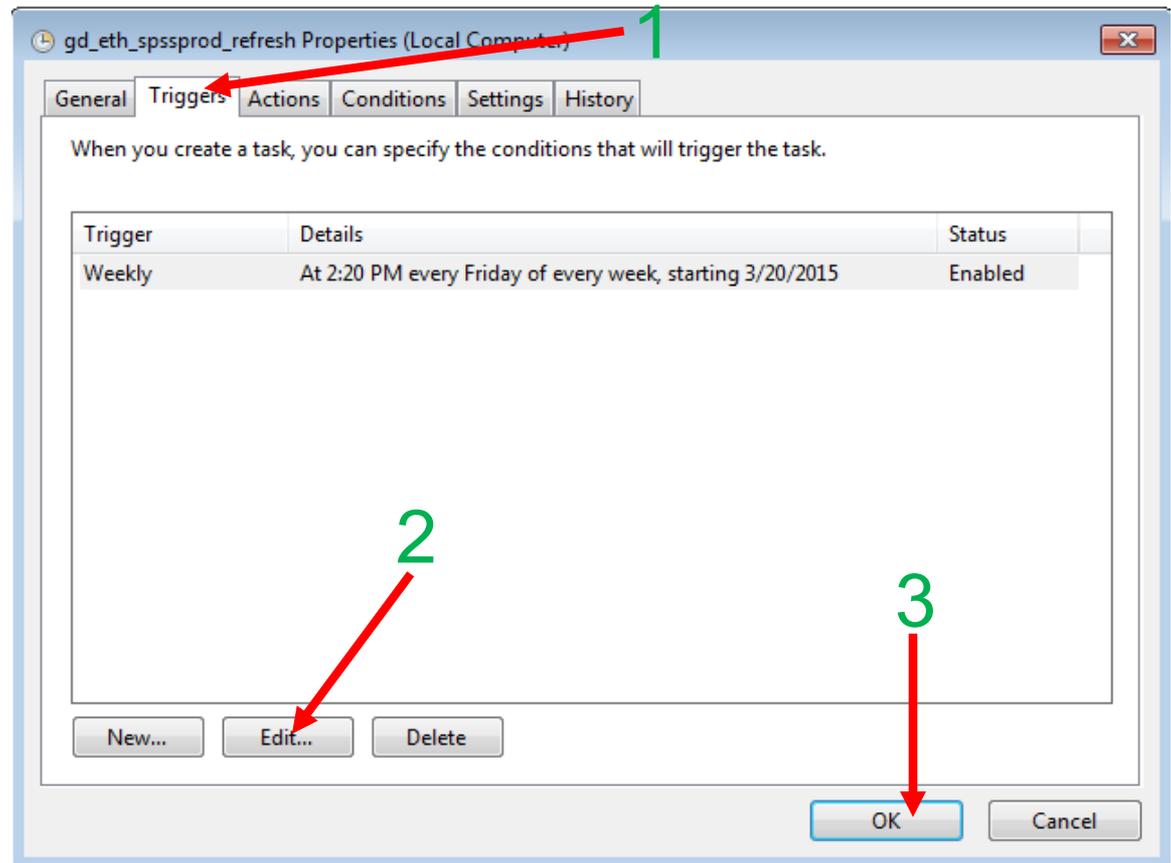


2. Click on Action on the menu bar
  - From here you can run, end, disable, delete and view the properties
3. Click on Properties to change the day, time or recurrence settings



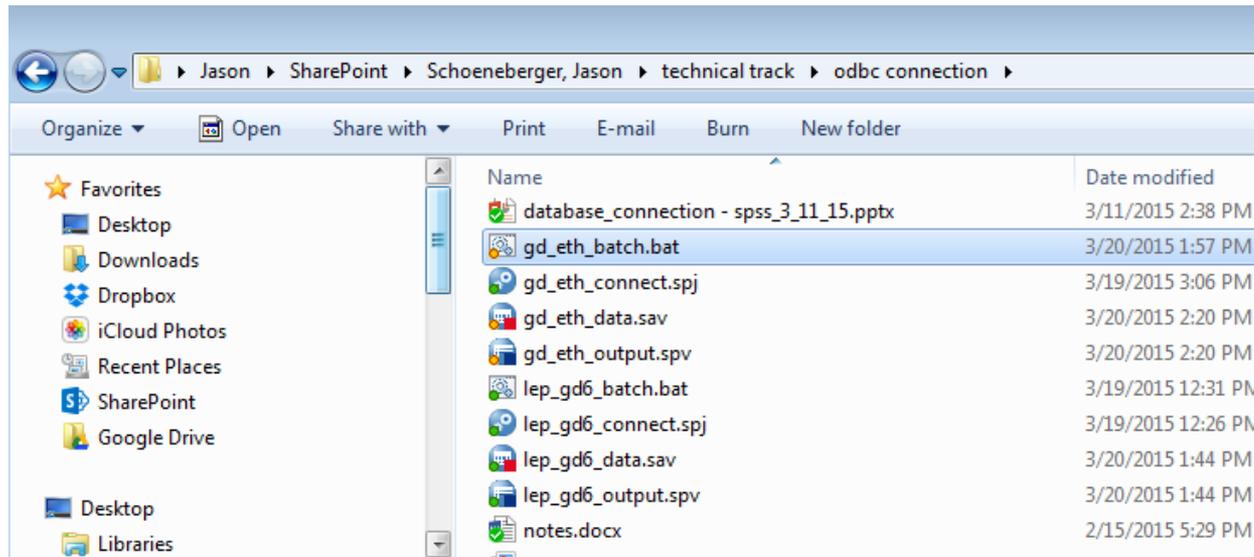
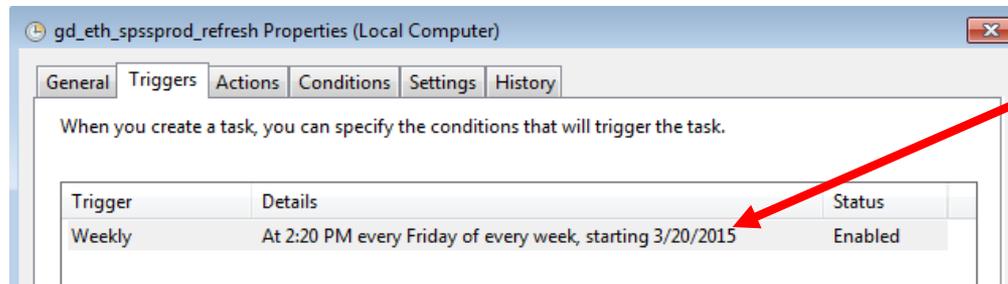
# Monitoring/editing a scheduled task

1. Using the tabs across the top of the Properties pane, you can edit various aspects of the scheduled task
2. On a particular tab, click Edit to change details about the scheduled task
3. Click OK when finished editing



# Alignment of task trigger and file time stamp

1. Note trigger of 2:20 PM on Fridays
2. Note time stamp on file when scheduled task runs



## Potential problems

---

- If time stamp on file fails to update...
  - On initial use, check code in .bat file to ensure no errors were made in copy-paste
  - Was your computer inadvertently shut-down or powered-down at the scheduled time?
  - Was the server or machine where database source is located inadvertently shut-down or power-down at the scheduled time?
  - Have there been any changes to the database source (e.g. Access SQL), such as table name changes, connection information change (path or server name)?

## Questions/Need help

---

### Contact:

Jason Schoeneberger, Ph.D.

Senior Researcher and Task Lead

REL Mid-Atlantic at ICF International

[jason.schoeneberger@icfi.com](mailto:jason.schoeneberger@icfi.com)

704-307-9395



Please visit [www.relmidatlantic.org](http://www.relmidatlantic.org) for other data tools!