

District Data Coordinator Toolbox: Automating Data Acquisition Using Database Connections in SAS

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.

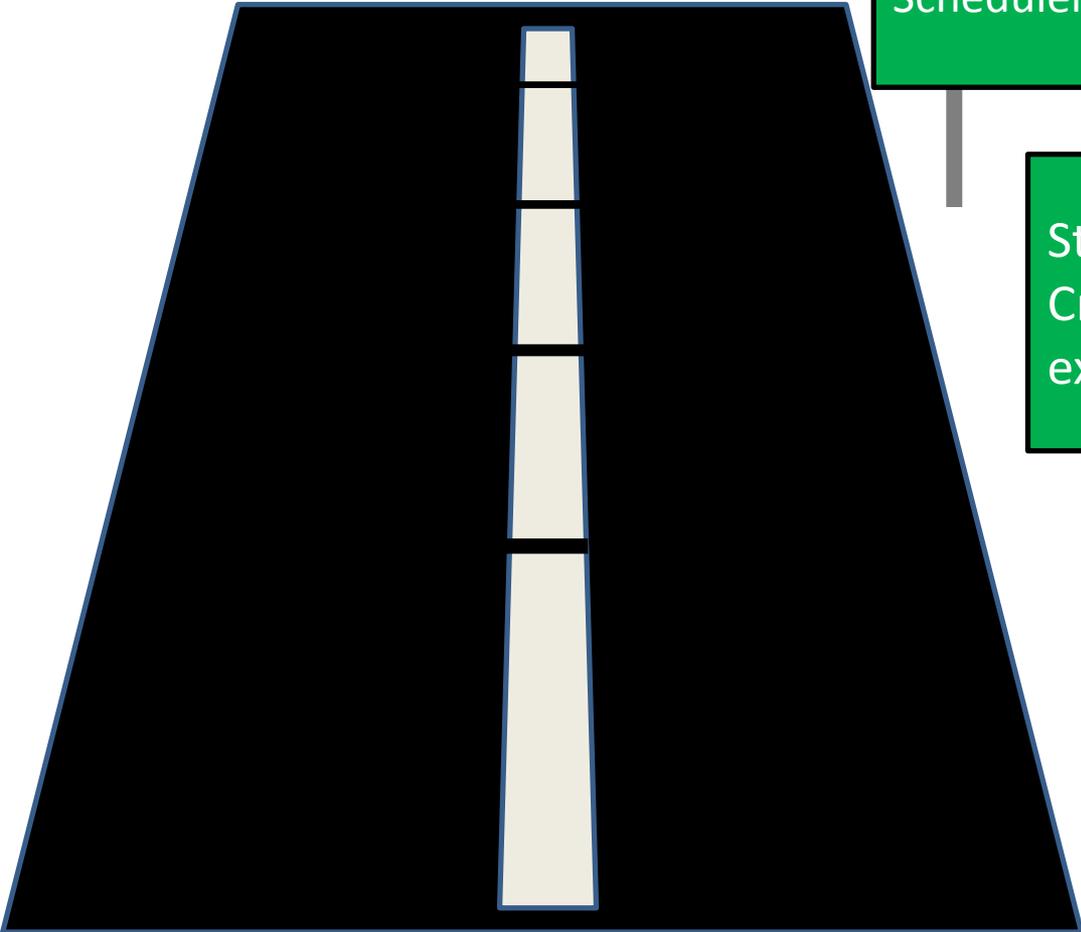
If not, please review the tool below, which is available on the REL Mid-Atlantic website:

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

Taking the next step...

- You're making use of a database connection(s) in SAS, 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 3.
Schedule task with Task Scheduler

Step 2.
Create batch file to execute SAS

Step 1.
Create SAS syntax file

Our road map to automation

1. We prepare an SAS syntax file to generate what we need
2. A batch file is created containing instructions to tell our computer to execute our SAS Program
3. 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 ethnicity	student_grade									All		
	06			07			08					
	n	row %	col %	n	row %	col %	n	row %	col %	n	row %	col %
African American	132	26.61	53.44	176	35.48	33.98	188	37.90	34.56	496	100.00	37.89
American Indian	1	20.00	0.40	2	40.00	0.39	2	40.00	0.37	5	100.00	0.38
Asian	7	21.21	2.83	13	39.39	2.51	13	39.39	2.39	33	100.00	2.52
Hispanic	54	27.69	21.86	74	37.95	14.29	67	34.36	12.32	195	100.00	14.90
Multi-Racial	6	24.00	2.43	13	52.00	2.51	6	24.00	1.10	25	100.00	1.91
White	47	8.47	19.03	240	43.24	46.33	268	48.29	49.26	555	100.00	42.40
All	247	18.87	100.00	518	39.57	100.00	544	41.56	100.00	1309	100.00	100.00

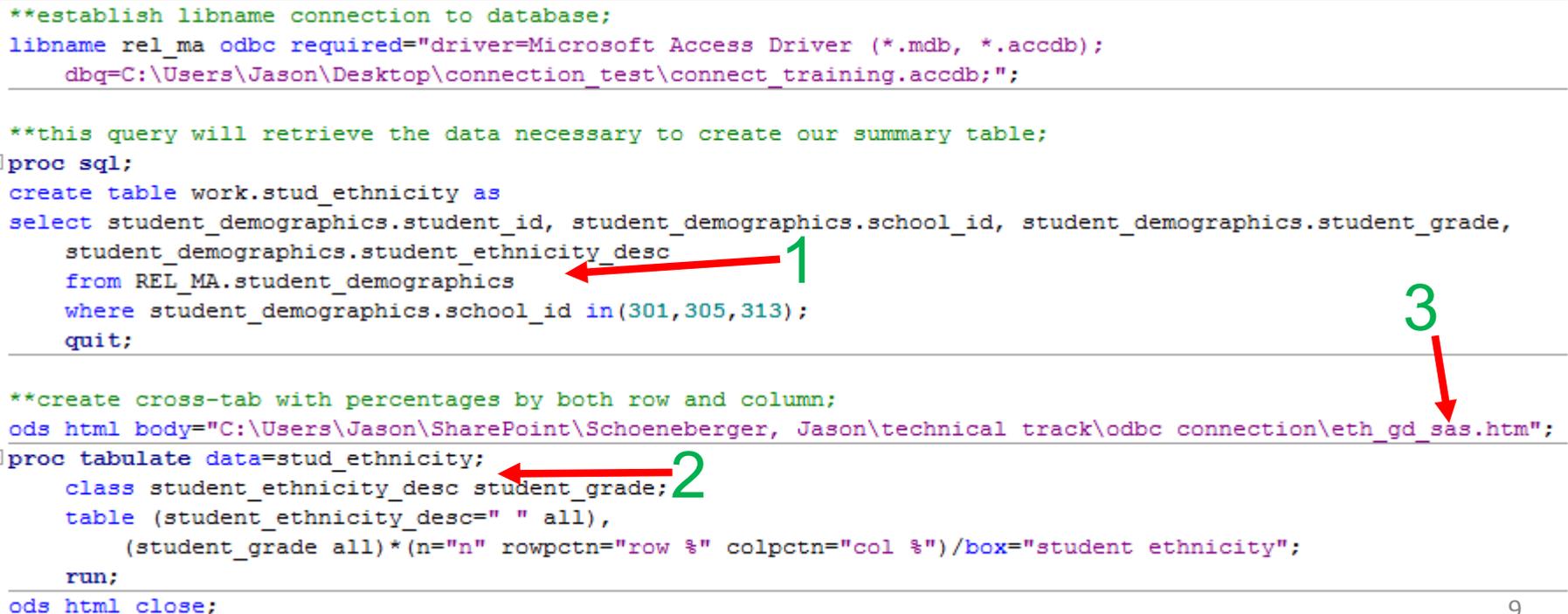
Using SAS syntax

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

Compiling SAS commands in syntax form

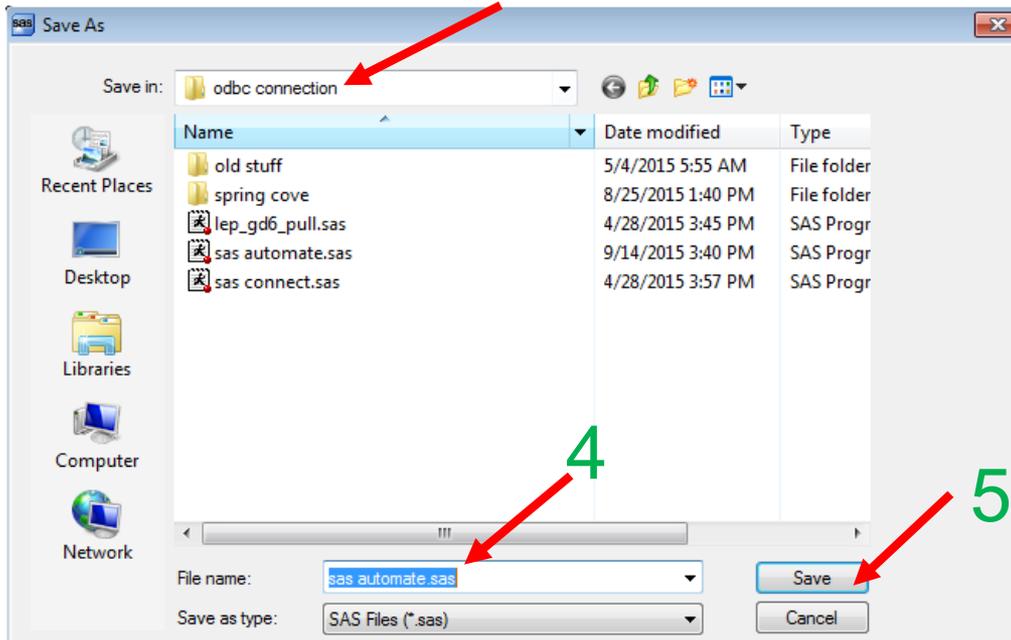
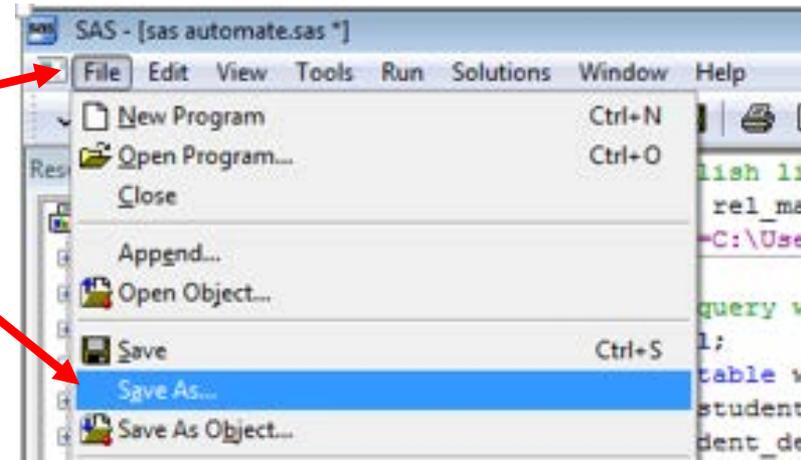
1. Obtain data to work with from database query
2. Analyze or manipulate data in logical order to generate desired results (here, a crosstab table)
3. Save output (as HTML output in desired path)

```
**establish libname connection to database;  
libname rel_ma odbc required="driver=Microsoft Access Driver (*.mdb, *.accdb);  
      dbq=C:\Users\Jason\Desktop\connection_test\connect_training.accdb";  
  
**this query will retrieve the data necessary to create our summary table;  
proc sql;  
  create table work.stud_ethnicity as  
  select student_demographics.student_id, student_demographics.school_id, student_demographics.student_grade,  
         student_demographics.student_ethnicity_desc  
  from REL_MA.student_demographics  
  where student_demographics.school_id in(301,305,313);  
quit;  
  
**create cross-tab with percentages by both row and column;  
ods html body="C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc connection\eth_gd_sas.htm";  
proc tabulate data=stud_ethnicity;  
  class student_ethnicity_desc student_grade;  
  table (student_ethnicity_desc=" " all),  
        (student_grade all)*(n="n" rowpctn="row %" colpctn="col %")/box="student ethnicity";  
run;  
ods html close;
```



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

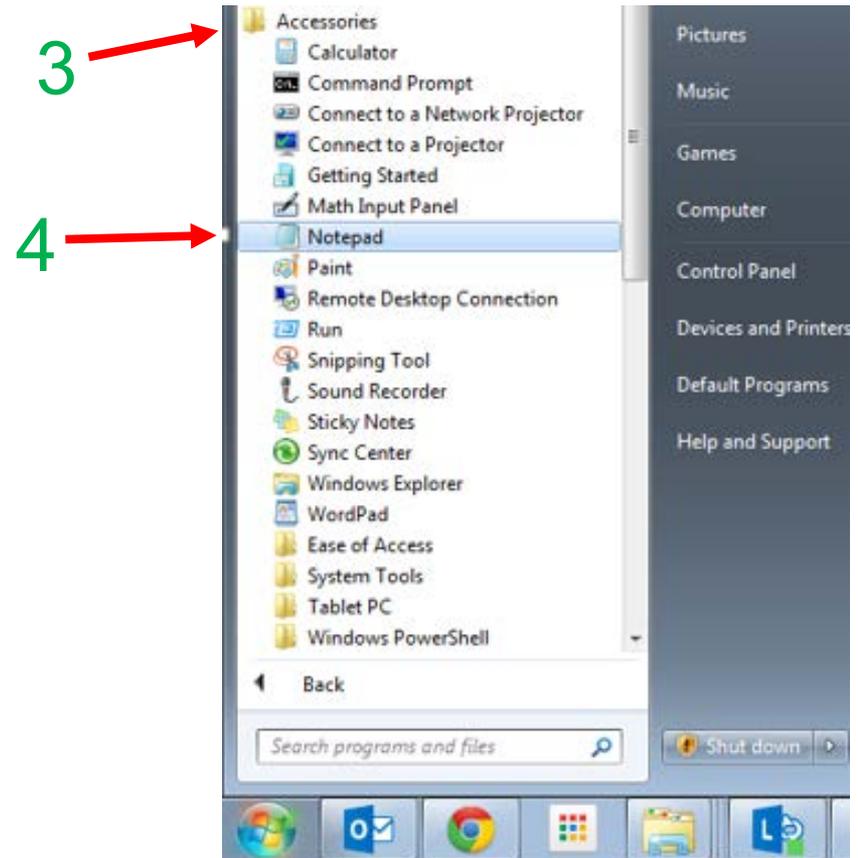
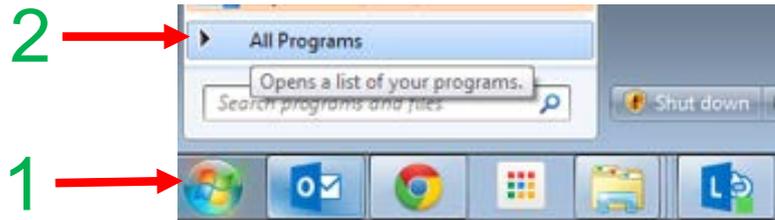


Creating a batch file

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

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
- All code should be on a **single** line

```
"c:\program files\SASHome\SASFoundation\9.4\sas.exe" -batch -sysin  
"C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc connection\sas  
automate.sas" -log "C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical  
track\odbc connection\sas_eth_grade.log" -print  
"C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc  
connection\sas_eth_grade.lst"
```

Creating the batch file

1. Specify the SAS executable path on your machine
2. Specify the path for the saved SAS program
3. Specify the paths for the SAS log and output files

The screenshot shows a Notepad window titled "gd_eth_batch.bat - Notepad" with a menu bar (File, Edit, Format, View, Help) and a text area containing the following batch script:

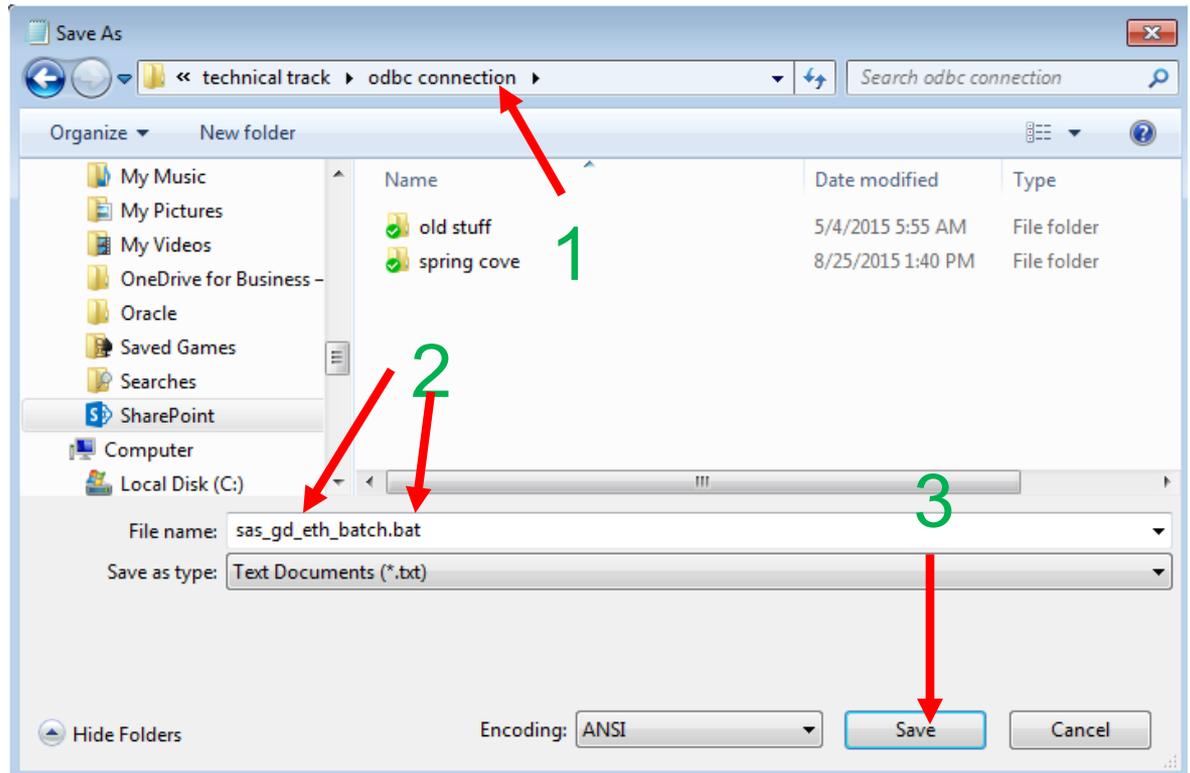
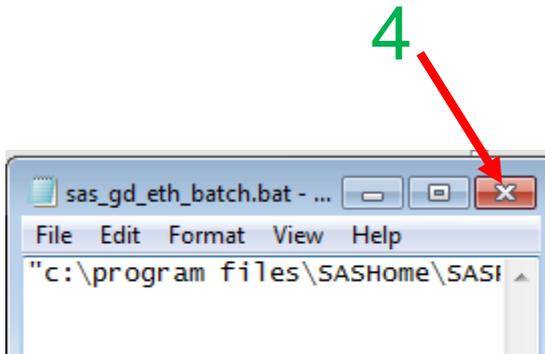
```
"c:\program files\SASHome\SASFoundation\9.4\sas.exe" -batch  
-SYSIN "C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc connection\sas_automate.sas"  
-LOG "C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc connection\sas_eth_grade.log"  
-PRINT "C:\Users\Jason\SharePoint\Schoeneberger, Jason\technical track\odbc connection\sas_eth_grade.lst"
```

Annotations 1, 2, and 3 are green numbers with red lines pointing to the first three lines of the script. Annotations 4 and 5 are green numbers with red arrows pointing to the "File" menu and the "Save As..." option in the menu, respectively.

4. Click File
5. Click Save As

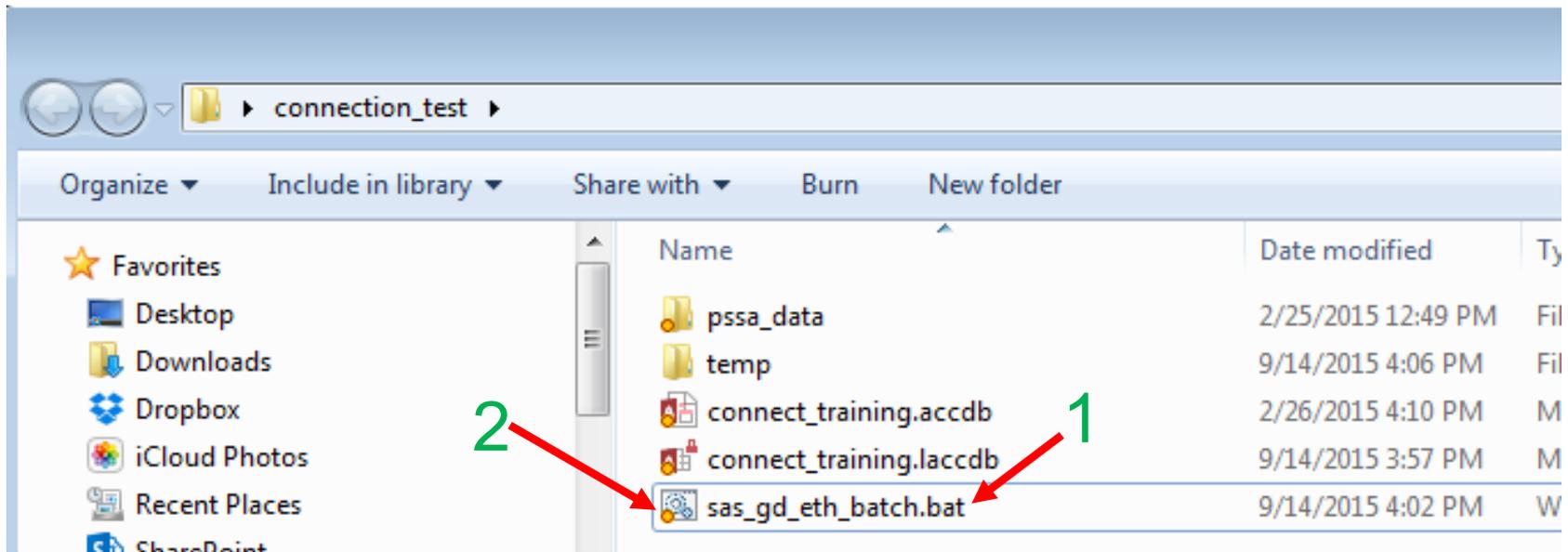
Saving the batch file

1. Navigate to your chosen folder
2. Name your batch file (sas_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 SAS open briefly, then immediately close
4. If there is an error in the code, you may get an error notice

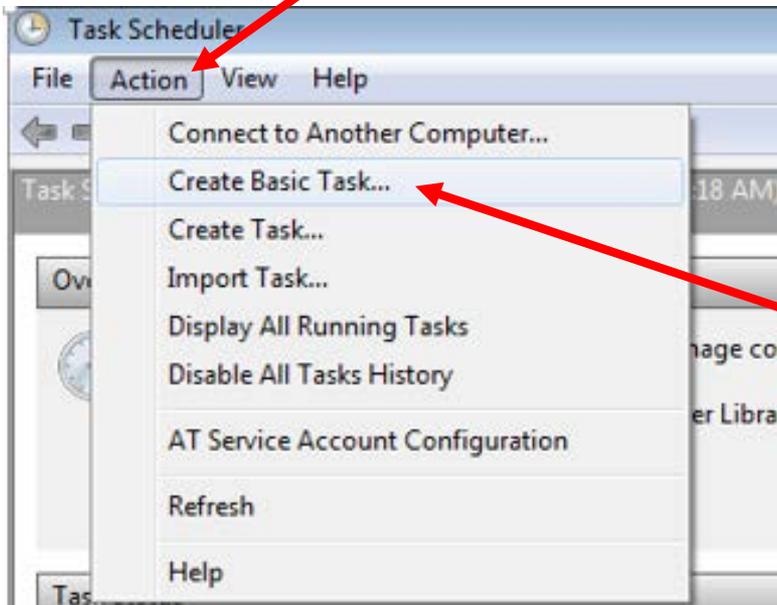
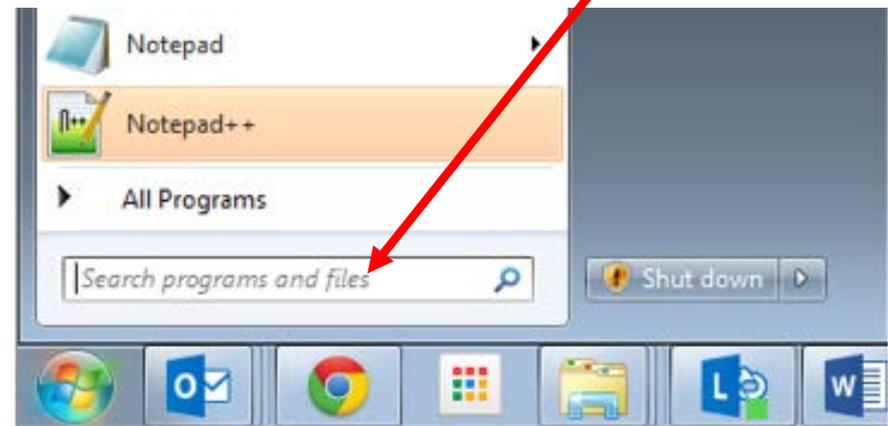


Scheduling a task

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

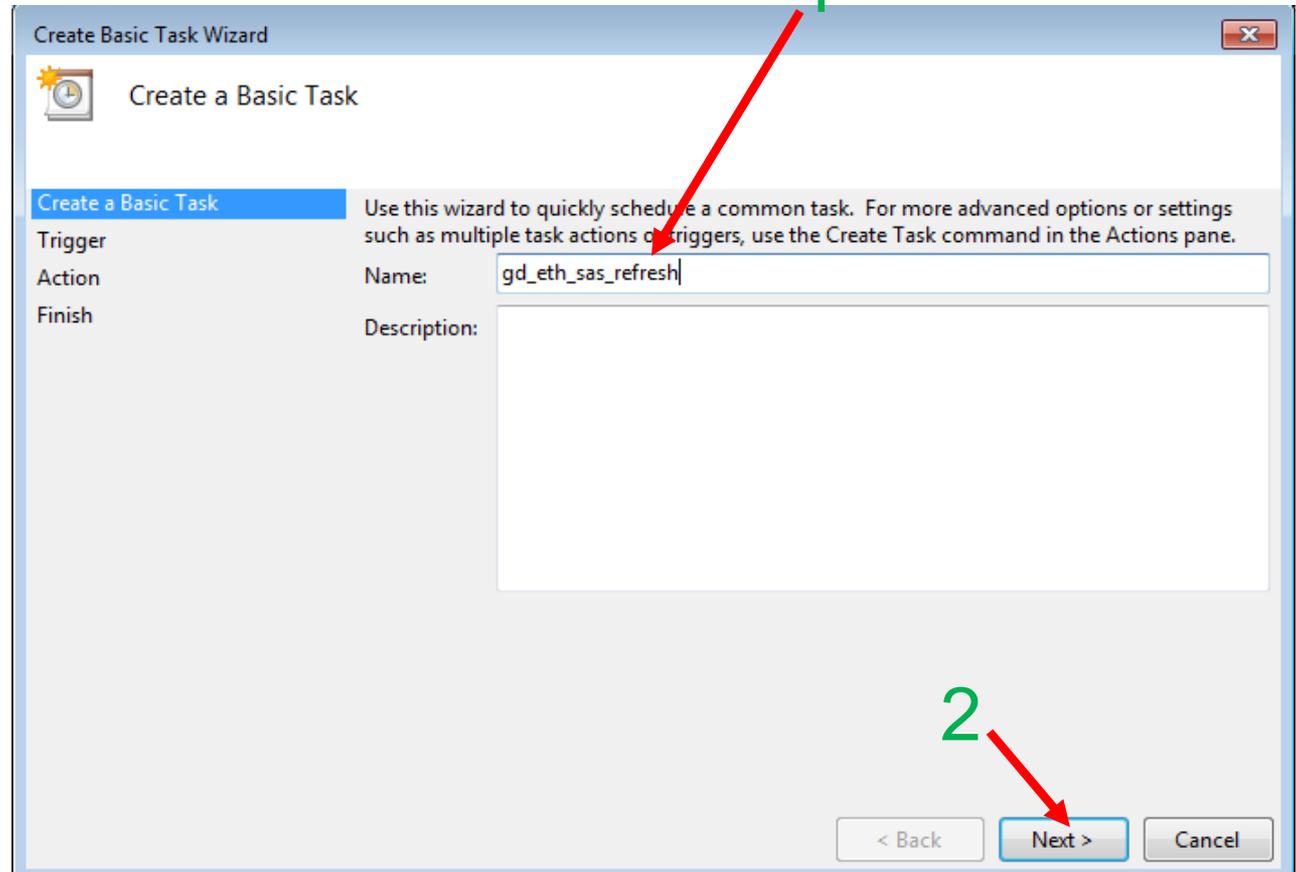
Opening Task Scheduler

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



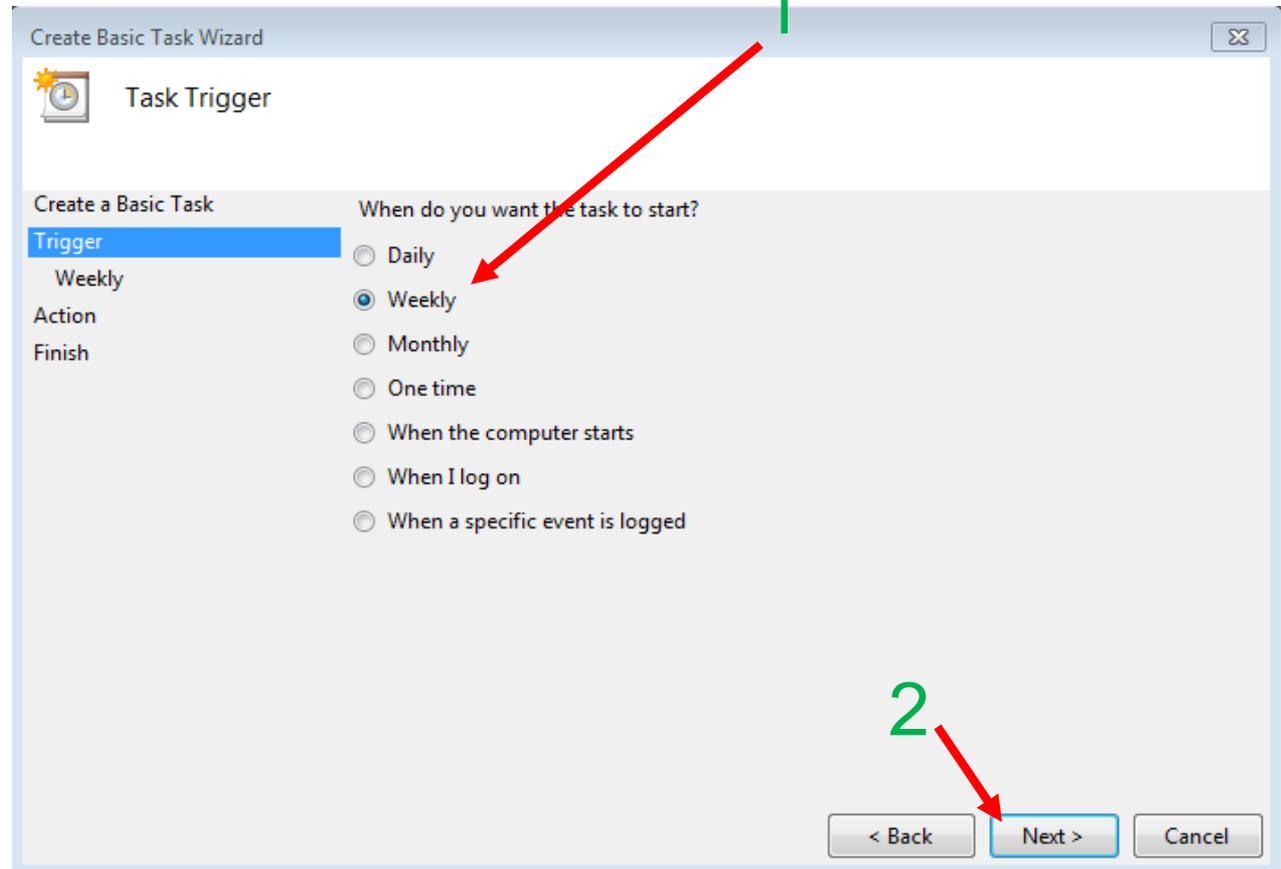
Naming the scheduled task

1. Enter a name for the scheduled task
(e.g gd_eth_sas_refresh)
2. Click Next



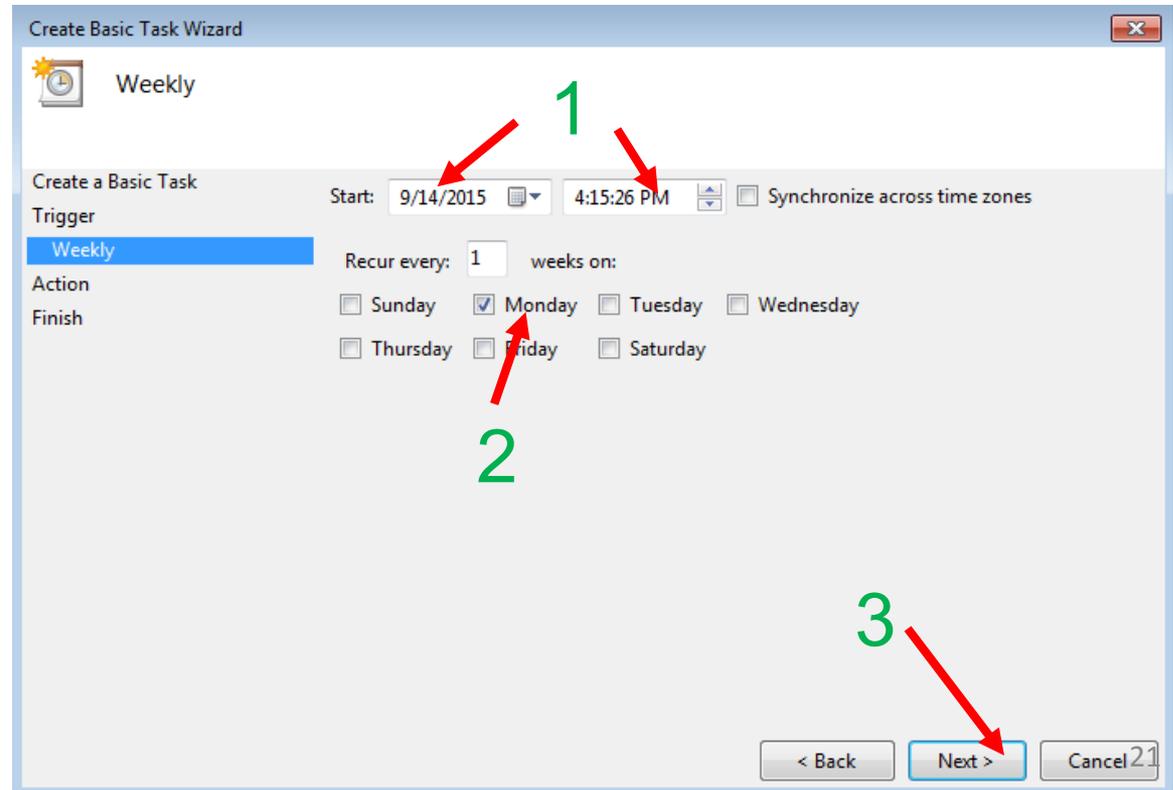
Set frequency of scheduled task

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



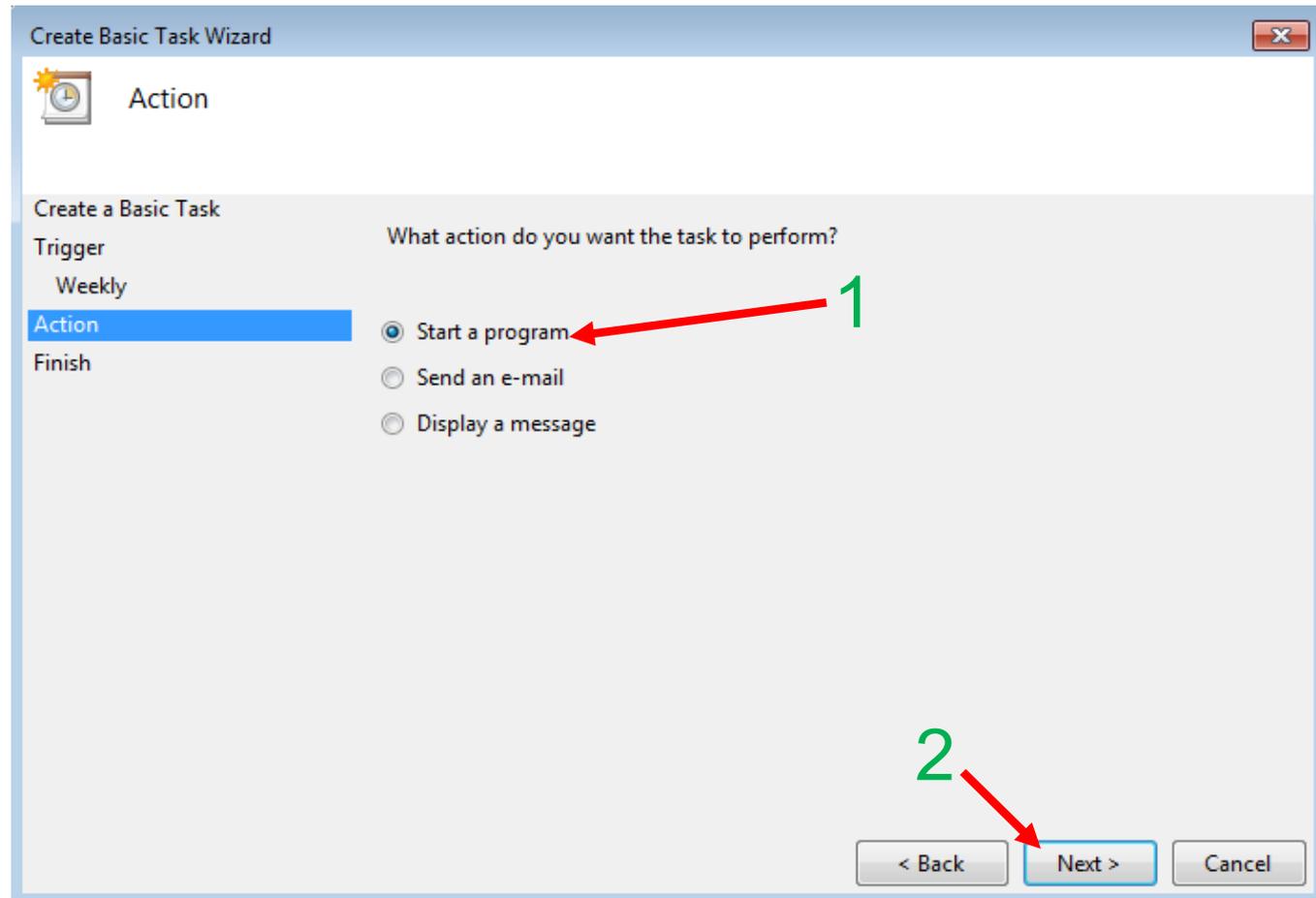
Set time-of-day and recurrence of scheduled task

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



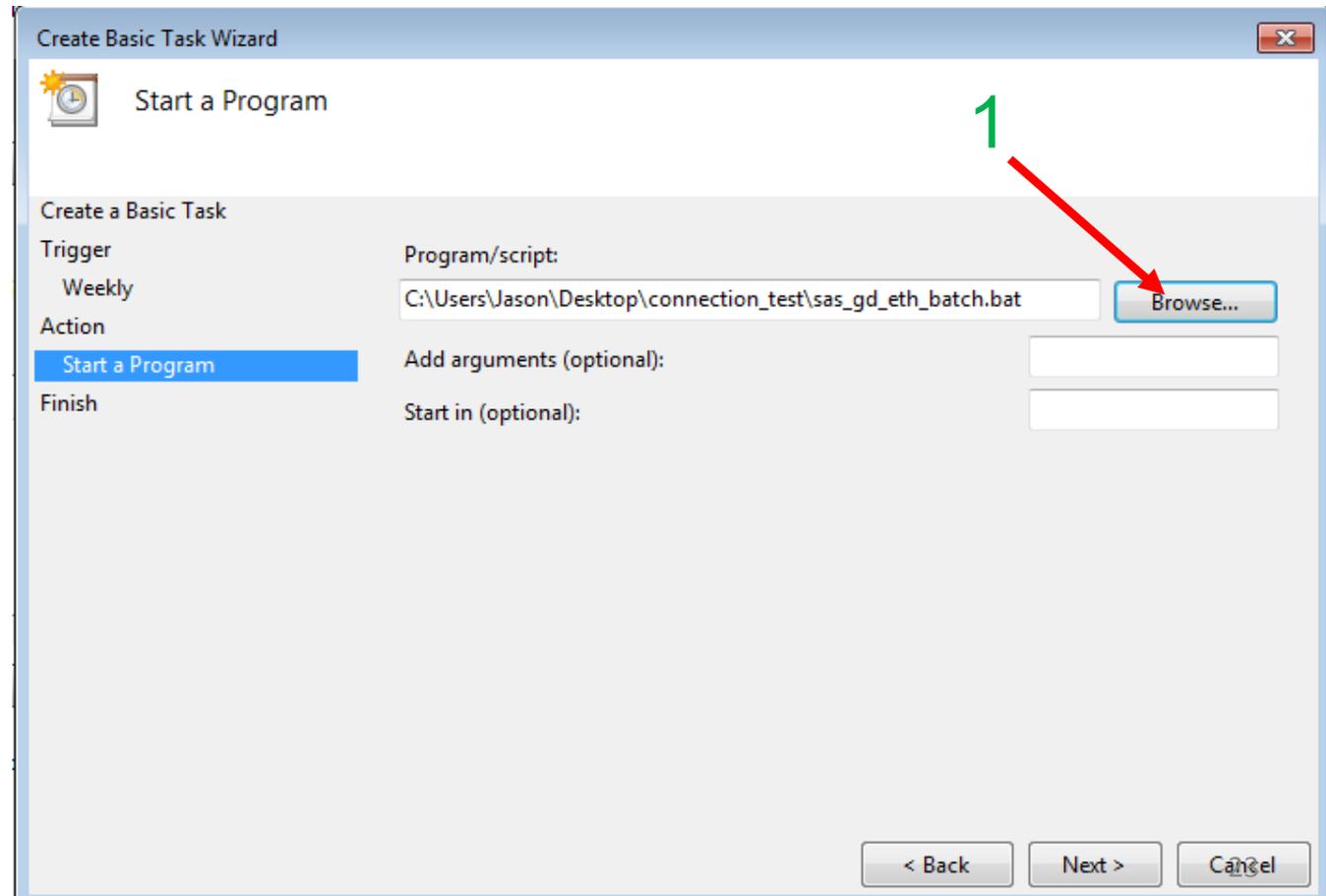
Specify the type of action to be scheduled

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



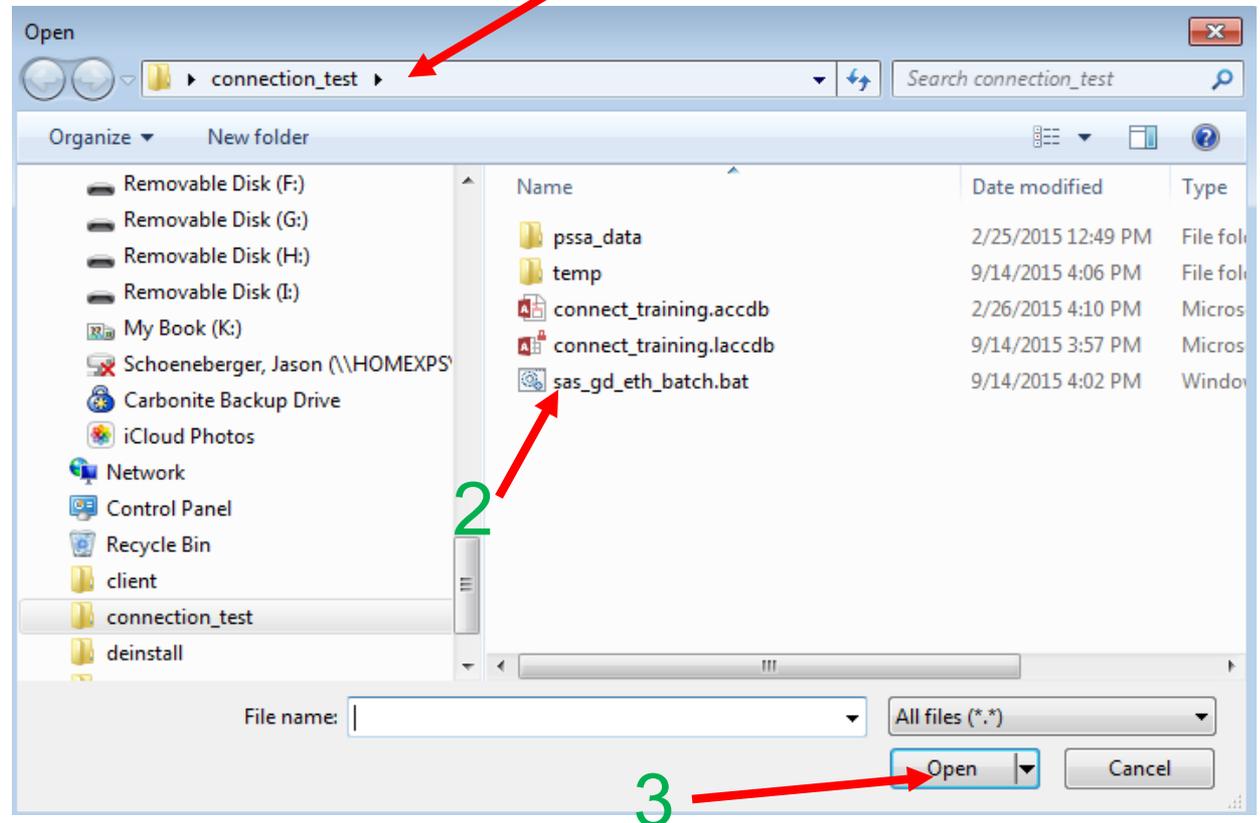
Specify file scheduled task should execute

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



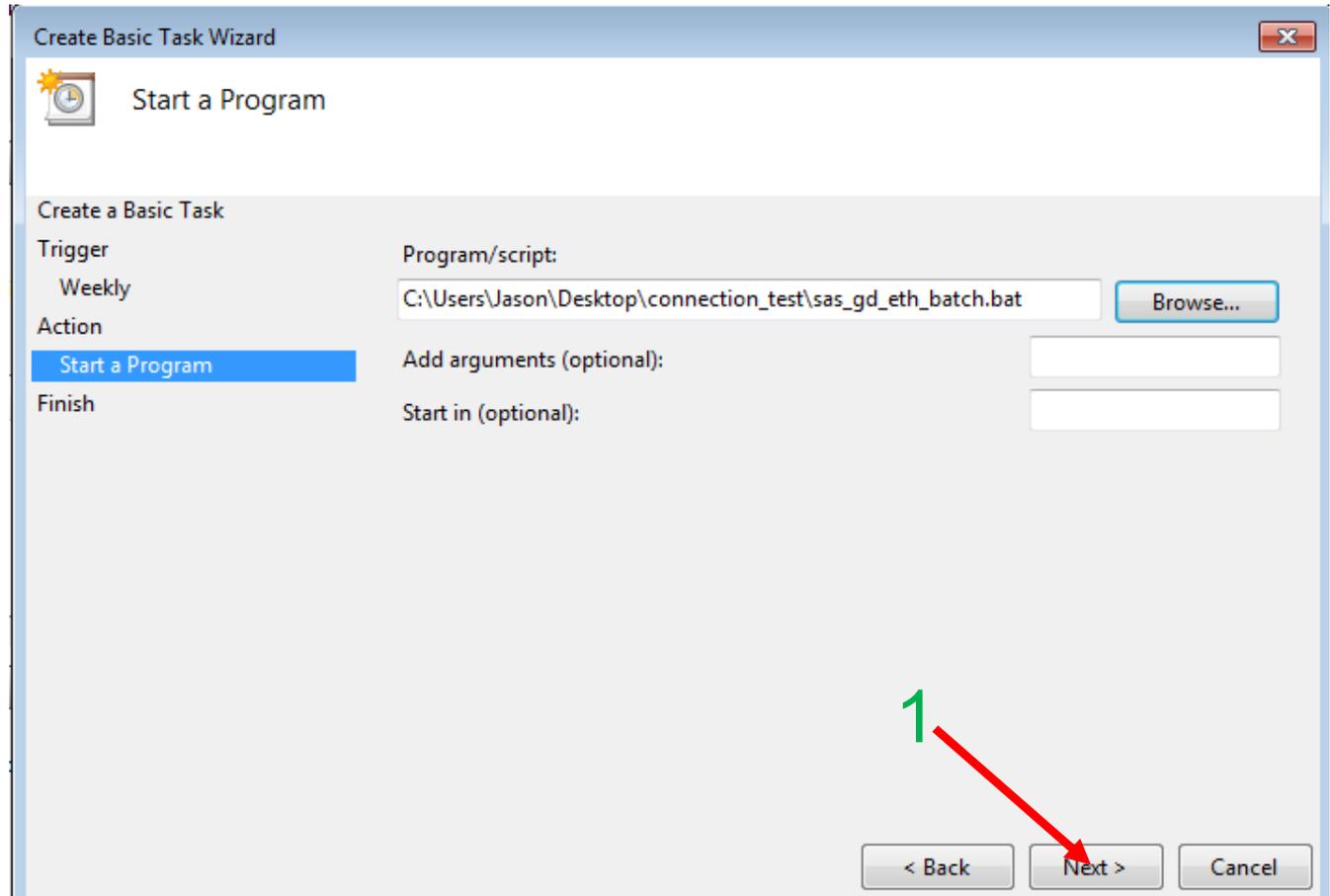
Specify batch file to execute

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

Create Basic Task Wizard

Summary

Create a Basic Task

Trigger

Weekly

Action

Start a Program

Finish

Name: gd_eth_sas_refresh

Description:

Trigger: Weekly; At 4:15 PM every Monday of every week, starting 9/14/2015

Action: Start a program; C:\Users\Jason\Desktop\connection_test\sas_gd_eth_batch.

Open the Properties dialog for this task when I click Finish

When you click Finish, the new task will be created and added to your Windows schedule.

< Back Finish Cancel

Monitoring/editing a scheduled task

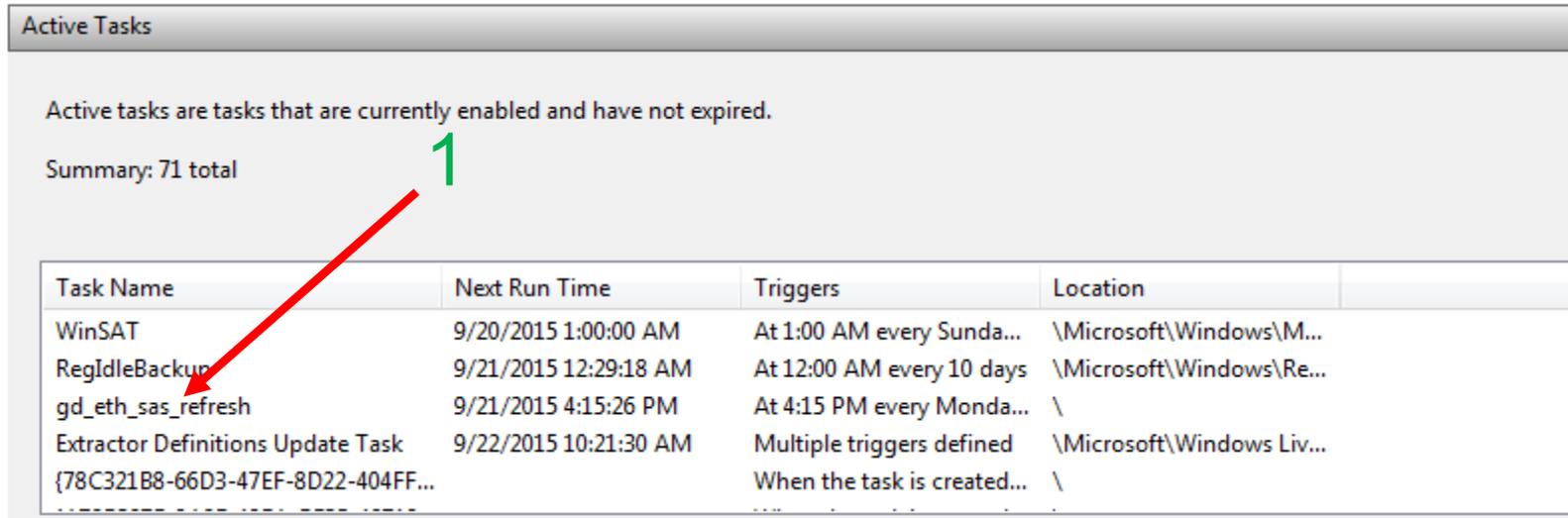
1. Double-click on the scheduled task under the Active Tasks pane inside Task Scheduler:

Active Tasks

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

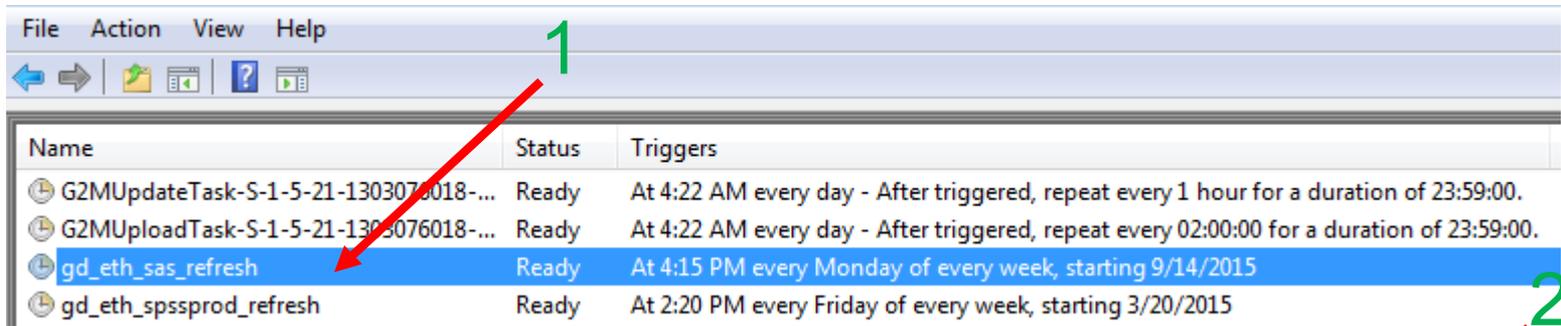
Summary: 71 total

Task Name	Next Run Time	Triggers	Location
WinSAT	9/20/2015 1:00:00 AM	At 1:00 AM every Sunda...	\Microsoft\Windows\M...
RegIdleBackup	9/21/2015 12:29:18 AM	At 12:00 AM every 10 days	\Microsoft\Windows\Re...
gd_eth_sas_refresh	9/21/2015 4:15:26 PM	At 4:15 PM every Monda...	\
Extractor Definitions Update Task {78C321B8-66D3-47EF-8D22-404FF...}	9/22/2015 10:21:30 AM	Multiple triggers defined When the task is created...	\Microsoft\Windows Liv...

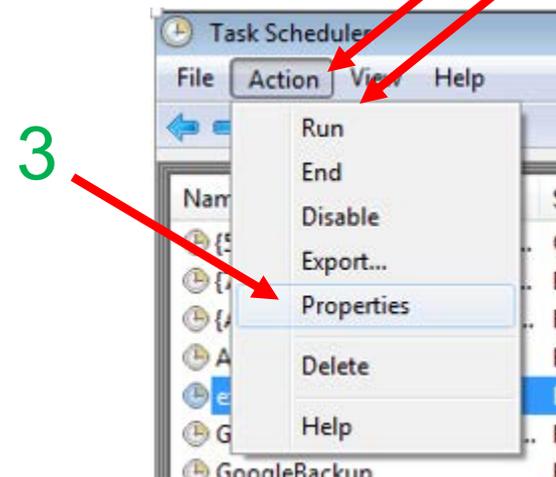


Monitoring/editing a scheduled task

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

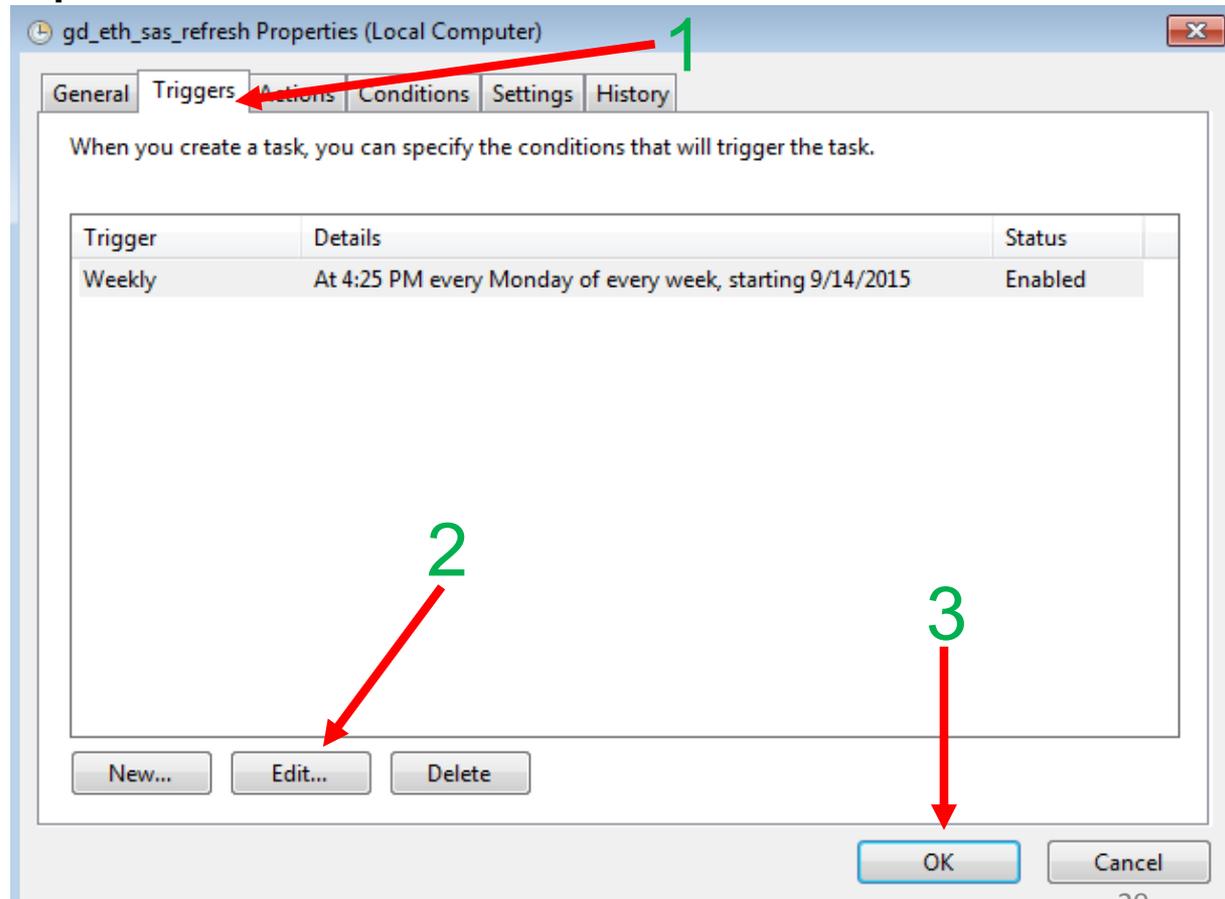


2. Click on Action on the menu bar
 - From here you can run, end, disable, delete and view the properties
3. Click on Properties to change the day, time or recurrence settings



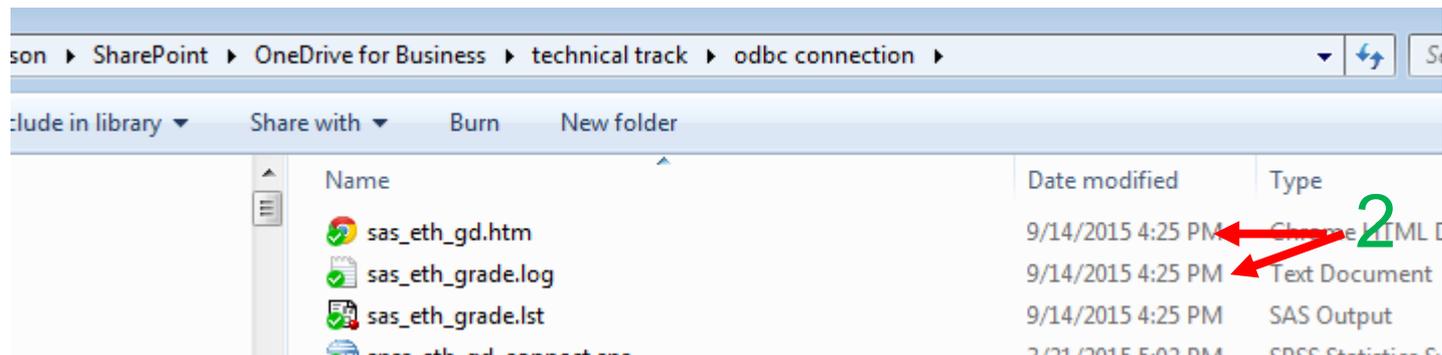
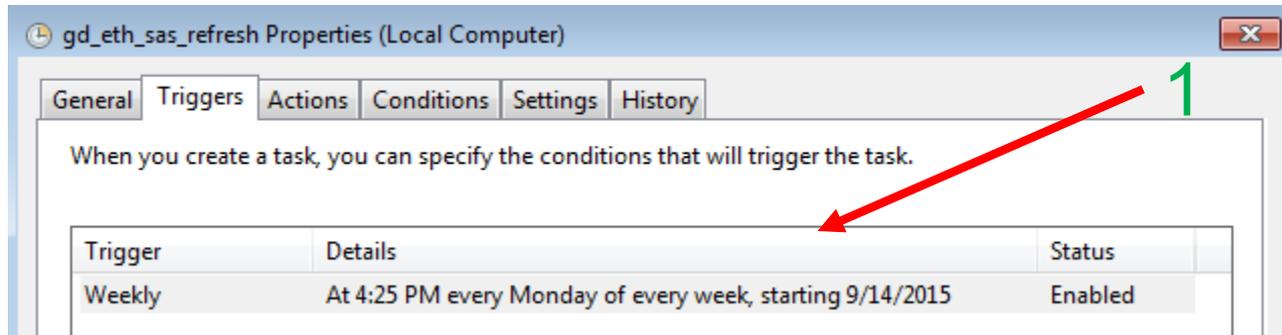
Monitoring/editing a scheduled task

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 4:25 PM on Mondays
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:

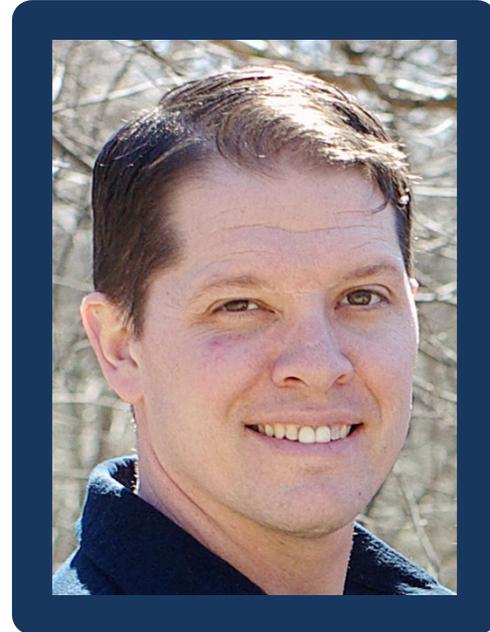
Jason Schoeneberger, Ph.D.

Senior Researcher and Task Lead

REL Mid-Atlantic at ICF International

jason.schoeneberger@icfi.com

704-307-9395



Please visit www.relmidatlantic.org for other data tools!