Fast-track SSIS

Sample manual - first two chapters



TABLE OF CONTENTS (1 of 7)

1	INTRODUCING SSIS	Page
1.1	Overview of SSIS	9
	SSIS Packages and Solution Explorer Control Flow	9 9
	Data Flow	10

USING VISUAL STUDIO	Page
Starting Visual Studio	11
Visual Studio and SQL Server Data Tools Running Visual Studio	11 11
Creating a Project	12
Visual Studio Windows	13
The Solution Explorer, Properties and SSIS Toolbox Windows	13
Floating Windows Auto-hiding Windows	14 14
	Starting Visual Studio Visual Studio and SQL Server Data Tools Running Visual Studio Creating a Project Visual Studio Windows The Solution Explorer, Properties and SSIS Toolbox Windows Floating Windows

3	PACKAGES	Page
3.1	Working with Packages	15
	Creating a Package	15
	Renaming Packages	15
	Opening or Editing Packages	16
	Viewing a Package's XML	16
	Saving a Package	17
	Closing Individual Packages	17
	Closing Multiple Packages	17
	Copying Packages	17
3.2	The Package Tabs	18
3.3	Running a Package	19
	Executing a Single Package	19
	Making a Package the Start-up Object and Debugging a Project	19
	Monitoring Package Execution	20
	Using the Progress Window	20
3.4	Stopping Running a Package	21

4	TASKS	Page
4.1	Working with Tasks	22
	Adding a Task	22
	Renaming a Task	22
	Editing a Task	23
	Executing an Individual Task	23
	Deleting and Suspending (Disabling) Tasks	24
4.2	Working with Task Diagrams	25
	Selecting Tasks	25
	Moving Tasks	25
	Auto-layout of Diagrams	26
	Zooming In and Out	26
4.3	Annotations	27
4.4	Grouping Tasks	28
	Grouping Tasks	28
	Grouping using Sequence Containers	29

5	DATA FLOW	Page
5.1	Introducing Data Flow	30
	Sources, Transforms and Destinations This and Subsequent Chapters	30 30
5.2	Data Flow Tasks	31
	Creating Data Flow Tasks Switching to Data Flow Tasks	31 31
5.3	Connections	32
	Types of Connection Project-Level and Package-Level Connections	32 32
	Creating Connections in Advance Creating Connections as Part of Sources/Destinations	33 33
5.4	Creating Our Package	34
5.5	The Advanced Editor	35

TABLE OF CONTENTS (2 of 7)

6	WORKING WITH SQL SERVER	Page
6.1	Creating a SQL Server Connection	36
6.2	Creating a SQL Server Source/Destination	37
	Step 1 – Creating the SQL Server Source/Destination	37
	Step 2 – Choosing the Table or Query	38
	Step 3 – Mapping Columns	38
6.3	Creating SQL Server Tables	39
6.4	Execute SQL Tasks	40
	Step 1 - Adding an Execute SQL Task	40
	Step 2 – Assign a Connection	40
	Step 3 – Enter a SQL Statement	41
	Step 4 – Optionally, Parse your SQL Command	41

7	WORKING WITH FLAT FILES	Page
7.1	Examples for this Chapter	42
7.2	Existing Flat File - Source	43
	Connecting to a Flat File Configuring Rows and Columns Creating the Flat File Source	43 44 44
7.3	Existing Flat File – Destination	45
7.4	Creating New Flat Files	45
	Step 1 – Creating a Destination Step 2 – Starting a New Connection Step 3 – Choose a File Format Step 4 – Browsing to a Folder Step 5 – Choosing a File Name Step 6 – Configuring the New File (Rows) Step 7 – Configure the New File (Columns) Step 8 – Configuring Mappings Step 9 – Running your Package	46 46 47 47 48 48 49 49
7.5	Using Multiple Flat Files Our Example	50 50
	Adding the MultiFlatFile Connector Using a MultiFlatFile Connector	51 52

8	WORKING WITH EXCEL	Page
8.1	Excel Workbooks – Overview	53
	The Example Used in this Chapter What Can Go Wrong	53 53
8.2	Importing/Exporting using Excel	54
	Step 1 – Creating an Excel Connection Step 2 – Creating a Source or Destination Step 3 – Mapping Columns	54 55 55
8.3	The 32-Bit Problem	56
	The Problem The Solution	56 57

9	DATA VIEWERS	Page
9.1	Using Data Viewers	58
9.2	Creating Data Viewers	59
	The Data Viewer Buffer	59

10	TRANSFORMS	Page
10.1	Adding Transforms	60
10.2	Sorting / Getting Unique Rows	61
10.3	Sampling	62
	The Sampling Transforms Incorporating Sampling Transforms into Data Flow	62 63
10.4	Combining and Splitting Data	64
	Introduction to the Transforms	64
	An Example using Multicast and Union All Transforms	65
	Configuring Union All Transforms	65
10.5	Aggregating Data	66

11	DATA TYPES	Page
11.1	The Data Types in SSIS	67
	The Three Data Type Families in SSIS Data Types Listed in this Chapter	67 67
11.2	SSIS Data Types	68
	String Data Types Integer Data Types Boolean or Logical Data Types Non-Integer Data Types Date/time Data Types Time Data Types	68 68 68 69 69

TABLE OF CONTENTS (3 of 7)

12	DATA CONVERSION TRANSFORMS	Page
12.1	Overview of Data Conversion	70
12.2	Creating Data Conversion Transforms	71
	Our Example: Importing Excel Unicode Data to Varchar Columns	71
	Adding a Data Conversion Transform	71
	Configuring a Data Conversion Transform	72
	Mapping Data Conversion Transform Columns	73
	The Complete Package	73

13	VARIABLES	Page
13.1	Overview of Variables	74
	bles	74
13.2	Working with Variables	75
	The Variables Window	75
	Creating a Variable	75
	Deleting Variables	75
	Changing the Scope of Variables	76
	Choosing What to See	76
13.3	User and System Variables	77

14	USING VARIABLES	Page
14.1	Example for this Chapter	78
	Creating the Variables Needed	78
14.2	Creating Row Count Transforms	79
	Adding a Row Count Transform	79
14.3	Creating Expression Tasks	80
14.4	Debugging	81
	Setting and Removing Breakpoints Debugging using Breakpoints	81 82

15	DERIVED COLUMN TRANSFORMS	Page
15.1	Overview of Derived Column Transforms	83
	Our Example – Calculating Film Length Statistics	83
	Adding a Derived Column Transform	83
15.2	Simple Column Expressions	84
15.3	Concatenating, Casting and Conditions	84
	Simple Concatenation Won't Work	85
	Casting	86
	The Ternary or Conditional Operator The Double = Sign When Testing	86 88
	Conditions	00
	The Relational Operators Allowed	88
15.4	More Examples and Syntax	89
	Example of a Mathematical Function	89
	Dealing with Nulls	90
	Example of a String Function	90

16	SSIS FUNCTIONS	Page
16.1	Working with Strings of Text	91
	Functions to Find and Replace Text Getting Extracts from a String of Text Other Text Functions New Line and Other Special Characters	91 91 92 92
16.2	Date Functions	93
	The DATEPART Function Using DATEADD to Add Dates Using DATEDIFF to Take the Difference between Two Dates	93 94 94
16.3	Casting Data	95

17	ERROR-HANDLING	Page
17.1	Our Example	96
17.2	Configuring Error-Handling	97
17.3	Tracking Error Rows	98

18	CONDITIONAL SPLITS	Page
18.1	Our Example	99
18.2	Creating a Conditional Split	100
	Beginning a Conditional Split The Operators Allowed Typing Conditions Spotting Errors (Red Text)	100 101 101 101
18.3	Directing Conditional Split Output	102

TABLE OF CONTENTS (4 of 7)

19	LOOKUP TRANSFORMS	Page
19.1	Our Example	103
19.2	Creating a Lookup Transform	104
	Step 1 – Creating the Transform Step 2 – Choose the Lookup Table Step 3 – Redirect Non-Matching Rows Step 4 – Choosing a Cache Mode Step 5 – Choosing Columns to Output Step 6 – Choosing Columns to Capture	104 104 105 106 107
19.3	Dealing with Matched Rows	108
19.4	Dealing with Unmatched Rows	109
	Strategy 1 – Store Unmatched Rows for Inspection Strategy 2 – Add Unmatched Shop Names as Unknown Strategy 3 – Add Unmatched Rows to the Lookup Table	109 110 111
19.5	Caching Lookup Tables	112
	Benefits of using Cache Transforms Our Example – Looking Up Property Types	112 112
	Filling the Cache Using the Cache	113 114

20	FILE SYSTEM TASKS	Page
20.1	Adding File System Tasks	115
20.2	Working with Files	116
	Creating a File Connection	116
20.3	Working with Folders	117
	Creating a Folder Connection Case Study: Creating a Folder	117 118

21	PRECEDENCE CONSTRAINTS	Page
21.1	Simple Constraints	119
21.2	Expression Constraints	120
	Automatically Annotating Expression Constraints	121
21.3	Combining Constraints	122

22	LOOPING OVER FILES	Page
22.1	Our Example	123
22.2	Setting up the Package	124
	Emptying the SQL Server Table Creating a String Variable to Hold Each File	124 124
22.3	Looping Over Files	125
	Step 1 – Adding the Foreach Loop Container Task	125
	Step 2 – Configuring the Loop Task	126
	Step 3 – Capturing the File Path in a Variable	127
	Step 4 – Checking the Package Works	127
22.4	Omitting Certain Files	128
22.5	The Data Import Task	129
	The Steps to Follow	129
	Step 1 – Create an Excel Connection	130
	Step 2 – Make this Excel Connection Dynamic	131
	Step 3 – Initialise the File Path Variable Value	132
	Step 4 – Configure and Finish the Data Flow Task	132
22.6	Running the Package	133

TABLE OF CONTENTS (5 of 7)

23	LOOPING OVER ADO ROWS	Page
23.1	How ADO Enumerators Work	134
23.2	Generating Multiple Files from a Single Table	135
	The Details of this Example	135
	Summary of the Steps to Follow	135
	Step 1 – Creating the Variables	136
	Step 2 – Create an Execute SQL Task to get Unique Shop Names	136
	Step 3 – Store the Results Set in the Object Variable	137
	Step 4 – Create the ADO Enumerator Loop	137
	Step 5 – Configure the ADO Enumerator Loop	138
	Step 6 – Add a Fixed Data Flow Task	139
	Step 7 – Make the Flat File Connection Dynamic	140
	Step 8 – Run your Package	140
23.3	Importing a Workbook's Worksheets	141
	Creating the Variable	142
	Create the Outline Package	142
	The Foreach Loop Type – Creating a Connection	143
	Choosing to Return the Workbook's Worksheet Names	144
	Creating the Data Flow Task (Source)	144
	Adding a Data Flow Destination	145
	Running the Package	145

24	SQL PARAMETERS	Page
24.1	Passing Parameters by Position	146
	Our Example Creating the Variables Creating the Loop over Files Creating the Execute SQL Task using Parameters The Parameter Mapping Choices	146 146 147 148
24.2	Running the Package Passing Parameters by Name	149 150
	Creating the Stored Procedure Changing the Execute SQL Task	150 150

25	MERGE JOINS	Page
25.1	What Merge Joins Do	151
25.2	How Merge Joins Work	152
25.3	Creating the Package	153
	Step 1 – Loading the Data Step 2 – Sorting the Data Step 3 – Merging the Data Step 4 – Picking out the Differences Step 5 – Finishing the Package	153 153 154 155 155

26	DEPLOYMENT	Page
26.1	Overview of Deployment	156
26.2	Preparing to Deploy	157
	Creating a Catalog Creating Folders	157 157
26.3	Deploying Projects and Packages	158
	Starting Project Deployment Finishing Deployment Deploying Individual Packages	158 159 159

EXECUTING DEPLOYED PACKAGES	Page
Executing Reports from the SSMS Menu	160
Viewing Reports	161
Viewing Package Reports Viewing Catalog Reports	161 161
Executing Packages from SQL	162
Scheduling Packages	163
Step 1 - Run SQL Server Agent Step 2 - Add a Job Step 3 - Adding Steps to the Job Step 4 - Choosing a Schedule	163 164 164 165 165
	PACKAGES Executing Reports from the SSMS Menu Viewing Reports Viewing Package Reports Viewing Catalog Reports Executing Packages from SQL Scheduling Packages Step 1 - Run SQL Server Agent Step 2 - Add a Job Step 3 - Adding Steps to the Job

28	PARAMETERS	Page
28.1	Overview of Parameters	166
	Our Example Variables and Parameters Needed	166 166
28.2	Creating Parameters	167
	Creating Project-Level Parameters Creating Package-Level Parameters	167 167
28.3	Setting up the Example	168
	The Basic Control Flow Tasks Starting to Configure the Loop over Files Making the Loop Folder Dynamic The Data Flow Task Excel Source Finishing the Data Flow Task Running the Package	168 168 169 169 170 170
28.4	Deploying Packages using Parameters	171
	Deploying Package-Level Parameters Deploying Project-Level Parameters	171 171
28.5	Running Deployed Packages with Parameters	172

TABLE OF CONTENTS (6 of 7)

29	SCRIPT TASKS	Page
29.1	Overview	173
29.2	Adding a Script Task	174
	Step 1 – Creating the Script Task Step 2 – Choose a Language Step 3 – Choosing the Start Program Name Step 4 – Editing your Script Step 5 – Understanding (and Tidying Up) the Code Generated Step 6 – Writing your Program	174 175 175 175 176
	Step 7 – Correcting any Errors Step 8 – Saving and Running your Code	177 178
29.3	Passing Variables to Script Tasks	179
	Passing the Variables to your Script Referencing Variables in Script A Short-Cut for Experienced Programmers The Most Common Mistake with Variables	179 180 180 181
29.4	Debugging Script	182
	Setting and Removing Breakpoints Removing All Breakpoints Stepping through Code Displaying Variable Values	182 182 183 184

30	PROGRAMMING (VARIABLES)	Page
30.1	Why Use Variables in Script?	185
30.2	Declaring Variables	185
	Declaring Variables Creating Nullable Variables Using Modified Hungarian Notation Default Values for Variables Problems with Declaring Variables within Clauses	186 186 186 187 187
30.3	Setting Values in Variables	187
	Declaring Integer Variables and Adding/Subtracting Accumulating Text in String Variables	188 188
20.4		
30.4	Variable Data Types	188
	Mapping C# Data Types to the CLR Runtime A Lazy Person's Data Types Logical Values Integers Decimal (Floating Point) Numbers Strings and Text Dates and Times Objects	189 189 190 190 190 190 191
30.5	Converting Variables	191
	Conversion Using Convert.To ToString() – Special Case for String Conversions Casting Data Types	192 192 192
30.6	Notes on Working with Specific Data Types	193
	Working with Characters Working with Strings Escape Characters Verbatim Strings Splitting Strings Formatting Dates Working with Numbers – Possible Operations	193 193 194 194 194 195

31	CONDITIONS	Page
31.1	Using IF for Conditions	196
31.2	Operators	196
31.3	The SWITCH statement	197
	The Syntax of SWITCH Limitations of SWITCH	198 199
31.4	Ternary and Coalesce Operators	199
	The Conditional Operator The Null Coalesce Operator	200 200

TABLE OF CONTENTS (7 of 7)

32	LOOPS	Page
32.1	Looping in C#	201
	Looping a Given Number of Times Looping While a Condition is True (While/Do)	201 201
32.2	Breaking Out of Loops	202

33	SCRIPT COMPONENTS	Page
33.1	Our Example	204
33.2	Starting the Package	205
	Removing Old Rows Creating a Flat File Data Connection Creating a Data Flow to Import Data	205 205 206
33.3	Creating your Script Component	207
	Step 1 – Add the Component Step 2 – Set the Columns Feeding into the Script Component	207 208
	Step 3 – Set the Columns Coming out of the Script Component	209
	Step 4 – Starting to Create Script	210
	Step 5 – Writing the Script Itself	211
33.4	Finishing the Package	212
	The Conditional Split Task	212
	The Valid Data Destination	213
	The Invalid Data Destination	213

34	SCRIPTING FILES AND FOLDERS	Page
34.1	Solutions Covered	214
34.2	Checking a Folder Exists	215
	Step 1 – Create the Variables Step 2 – Create the Script Task Step 3 – Tidy your Code and Reference System.IO Step 4 – Get a Reference to the Folder Step 5 – Creating the Expression Constraints	215 215 216 216 217
34.3	Checking a File Exists	218
	Step 1 – Create Two Versions of the Source Workbook	218
	Step 2 – Create the Variables that you will Need	219
	Step 3 – Create a Static Package	219
	Step 4 – Create a Script Task	220
	Step 5 – Write your Script	220
	Step 6 – Create Two Expression Constraints	221

35	SENDING MAIL	Page
35.1	The Send Mail Task	222
35.2	Sending Mail in Script	223
	Creating and Passing Variable Values Referencing the Mail Namespace Writing your Script	223 224 224

36	EVENTS	Page
36.1	Overview of Events	225
	Seeing Events on the Progress Tab Our Example	225 225
36.2	The List of Events	226
36.3	Using Event-Handlers	227
	Creating (and Deleting) Event-Handlers The Event-Handlers for our Example	227 228

37	AUDIT TRANSFORMS	Page
37.1	Using Audit Transforms	229
	What Audit Transforms Do Adding Audit Transforms Choosing Columns to Include	229 229 230
	Running an Audit Transform	230

38	LOGGING	Page
38.1	Overview of Logging	231
	Showing the SYSSSISLOG Table in SQL Server	231
	A Typical Text Log	231
38.2	Starting to Log	232
	Beginning a Log	232
	Choosing a Logging Provider	232
	Specific Notes on Creating a SQL Server Log	233
	Specific Notes on Creating a Text File Log	233
38.3	Configuring your Log	234
	Choosing which Tasks to Log	234
	Choosing which Events to Log	234
38.4	Catalog Logging	235

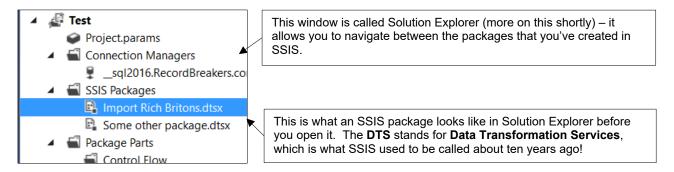
CHAPTER 1 - INTRODUCING SSIS

1.1 Overview of SSIS

SQL Server Integration Services (SSIS) is a program which allows you to build packages to Extract, Transform and Load data (it's often called an ETL application for this reason).

SSIS Packages and Solution Explorer

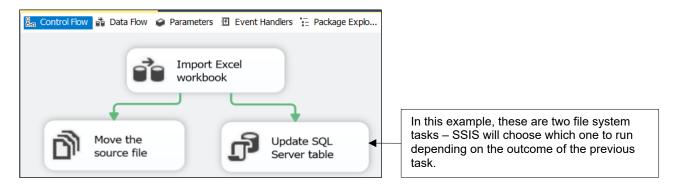
A package in SSIS is a file with the extension .dtsx:



A package consists of two main parts: control flow and data flow, as shown under separate headings below and overleaf.

Control Flow

The *control flow* part of a package consists of a series of instructions you ask SSIS to execute:





When you tell your child to tidy their room, do their homework and then come and help make dinner, this is a series of control flow tasks (as well as being a tad optimistic).

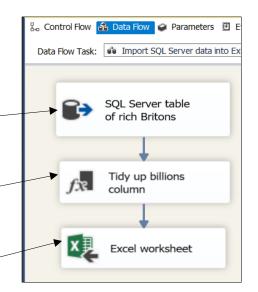
Data Flow

You can open any single data flow task to reveal instructions on how to load, transform and store data:

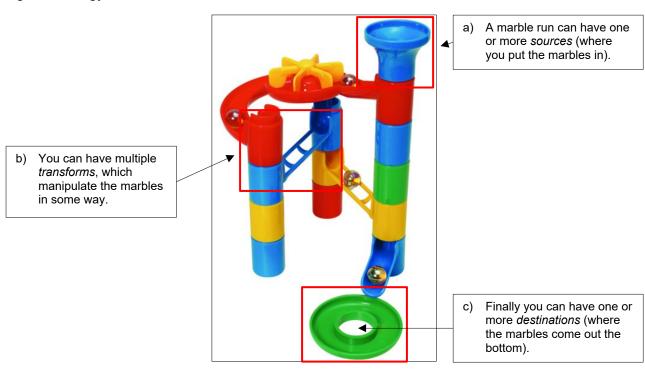
This is a *source* (where the data comes from – in this case a SQL Server table).

This is a *transform* (in this case, it's tidying up a numerical column).

This is a *destination* (where the data ends up – in this case in an Excel workbook).



A good analogy for data flow is a marble run:





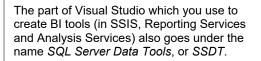
At a rough guess, you typically will spend about 70% of your time in Integration Services inside data flow tasks, and 30% in control flow.

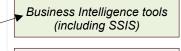
CHAPTER 2 - USING VISUAL STUDIO

2.1 Starting Visual Studio

Visual Studio and SQL Server Data Tools

Visual Studio is Microsoft's flagship development application – you can use it to create websites, mobile phone apps, SSIS packages and Windows applications, among other things:





Windows applications

(Visual Basic and C#)

ASP.NET (websites)

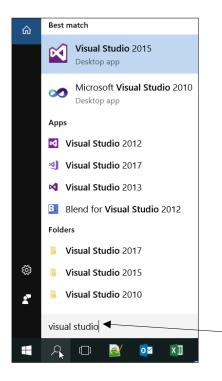
Mobile phone apps (Apple and Android)



What the above shows is that it doesn't matter whether you run Visual Studio or SQL Server Data Tools – the second is just a subset of the first.

Running Visual Studio

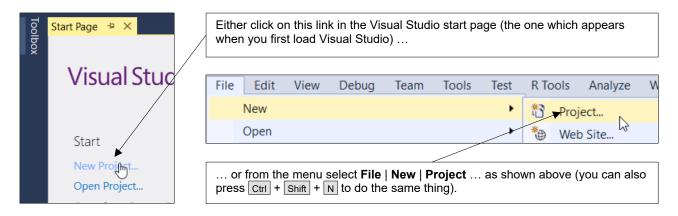
You can start Visual Studio in many ways – one method for Windows 10 is shown here:



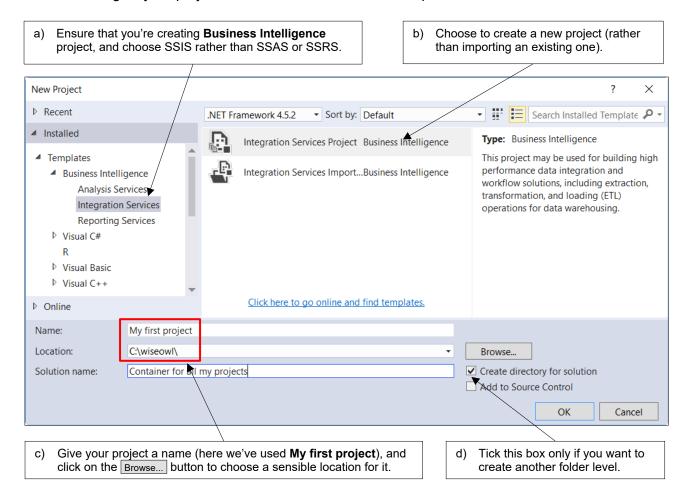
Click on the magnifying glass icon, then type in the program you want to run. This Wise Owl has got four different versions of Visual Studio on his computer! The one used in writing this manual was Visual Studio 2015.

2.2 Creating a Project

You store SSIS packages in a container called a *project*. Here's how to create one:



You can now give your project a name, and choose where to put it:



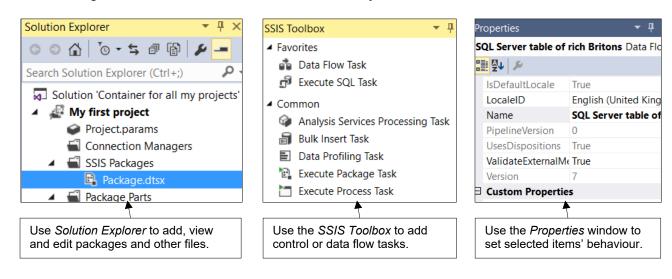
If you choose to create a directory for your solution as above, you'll end up with a long path!



2.3 Visual Studio Windows

The Solution Explorer, Properties and SSIS Toolbox Windows

When using SSIS, there are three main windows that you will use:



Here's how to display these three windows:

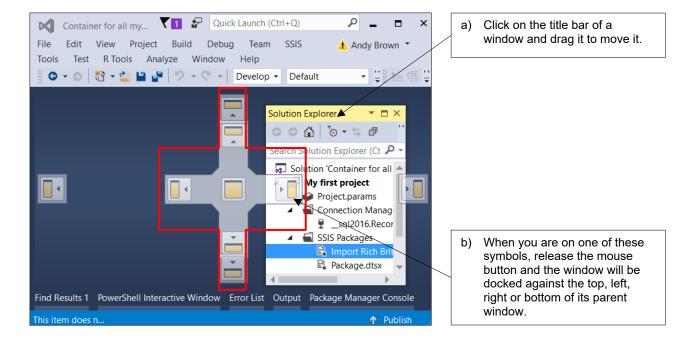
Window	Top menu	Keyboard	Other method
Solution Explorer	View → Solution Explorer	Ctrl + Alt + L	
SSIS Toolbox	SSIS → SSIS Toolbox		SSIS Toolbox
Properties	View → Properties Windows	F4	



To get the SSIS toolbox menu to appear you often have to open a package and click on it first. Don't confuse the (invaluable) SSIS toolbox with the (useless, in this context) standard Visual Studio toolbox.

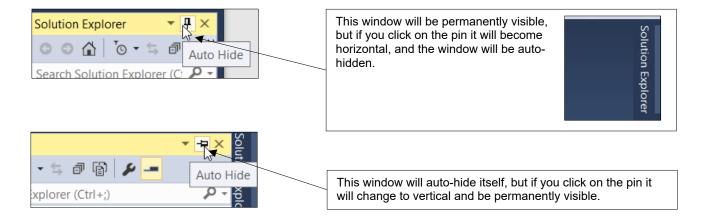
Floating Windows

You can click and drag on the title bar of any window to reposition it:



Auto-hiding Windows

You can click on the pin at the top right of any docked window to make it shrink when you're not using it:



What we do!

		Basic training	Advanced training	Systems / consultancy
		g		oonoana,
	Microsoft Excel			
Office	VBA macros		00	00
Off	Office Scripts			
	Microsoft Access	00	00	
Business ntelligence	Power BI	00	00	
	Power Apps			
Bus	Power Automate / PAD	00		
	SQL			
'er	Reporting Services		00	00
Ser	Report Builder			
SQL Server	Integration Services			
	Analysis Services			
Coding	Visual C# programming	1		
	VB programming			00
	DAX		00	00
	Python	0		



