

Building an Azure Business Intelligence Solution End to End

Hands On Workshop

Paul Andrew | Senior Consultant

Terry McCann | Principal Consultant

Simon Whiteley | Cloud Architect



Gold Data Analytics
Gold Data Platform
Gold Cloud Platform



<https://github.com/Adatis>

ModernDataWarehouseWorkshop

Agenda for the Day

Module 1

Microsoft Azure

Module 2

Storage
Uploading Data
Data Lake

Module 3

Real-time Data
Streaming
Power BI

Module 4

U-SQL - Data
Transformation
Basics

Module 5

USQL - Advanced
Analytics
Cognitive Services

Module 6

Data Factory
Orchestration
Dynamic Pipelines

Module 7

Data Presentation
& Consumption
Power BI Models

Module 8

Other Services
Q&A

Module Agenda

What is U-SQL

Extract
Select
Output

Logical Instance

Table
Views
Stored Procedures

Module Agenda

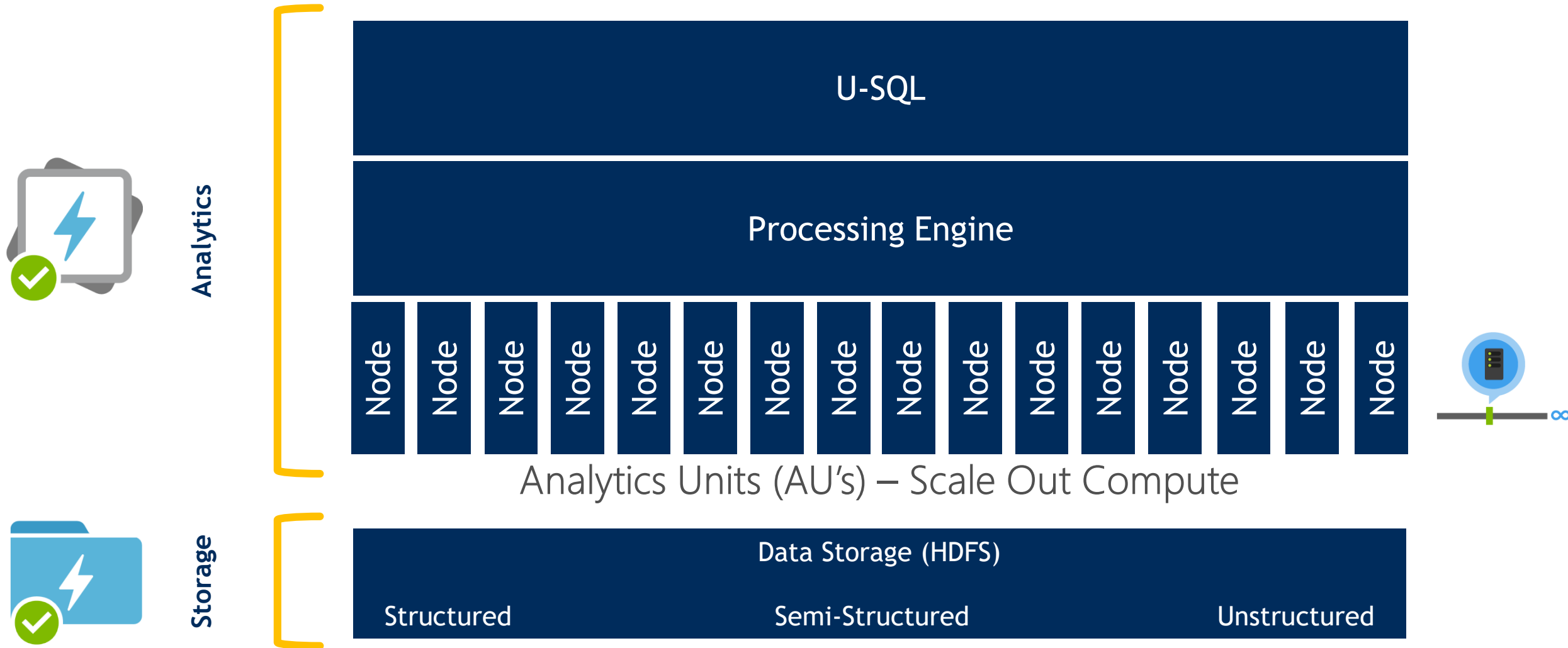
What is U-SQL

Extract
Select
Output

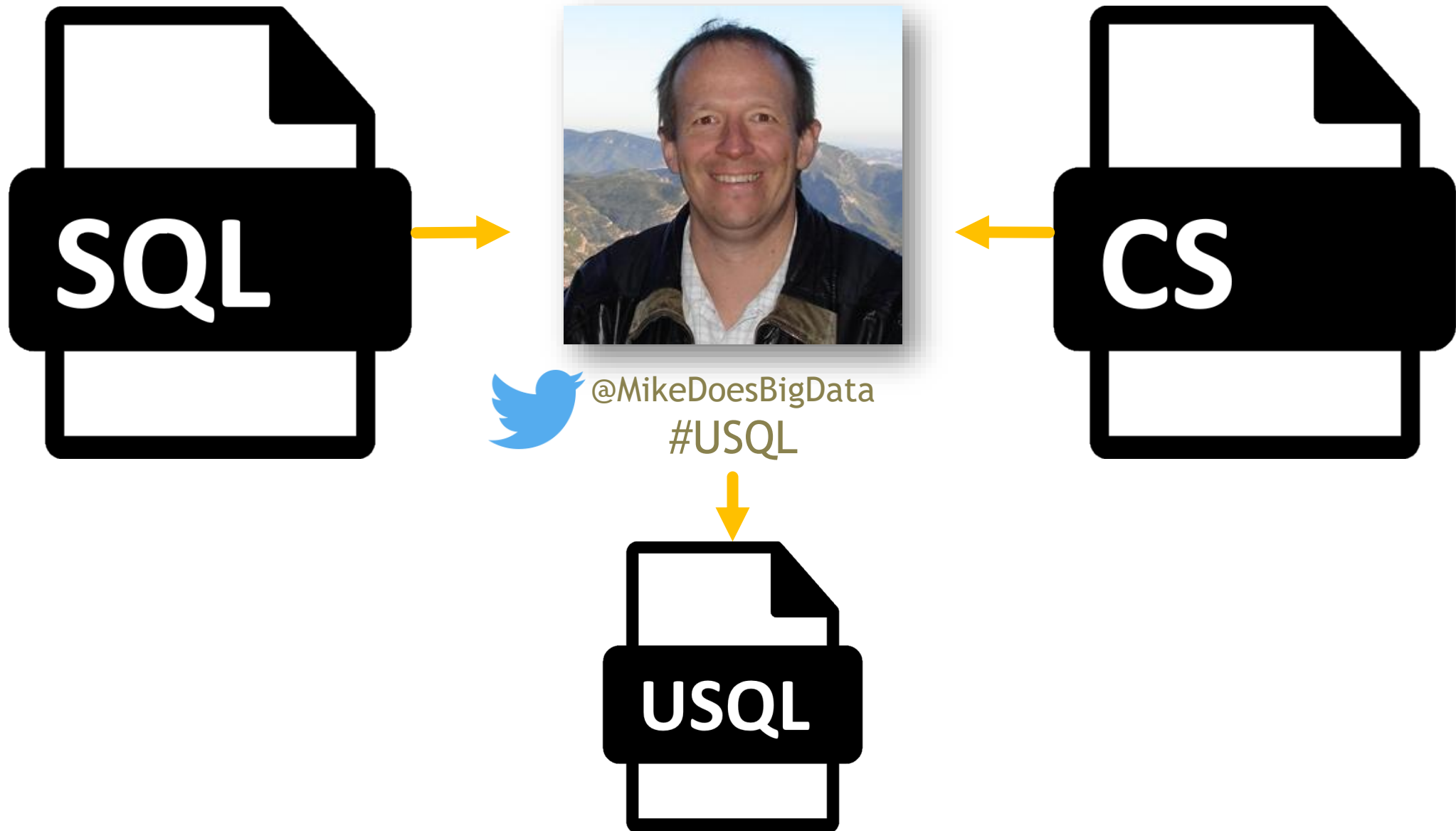
Logical Instance

Table
Views
Stored Procedures

Azure Data Lake – A Recap

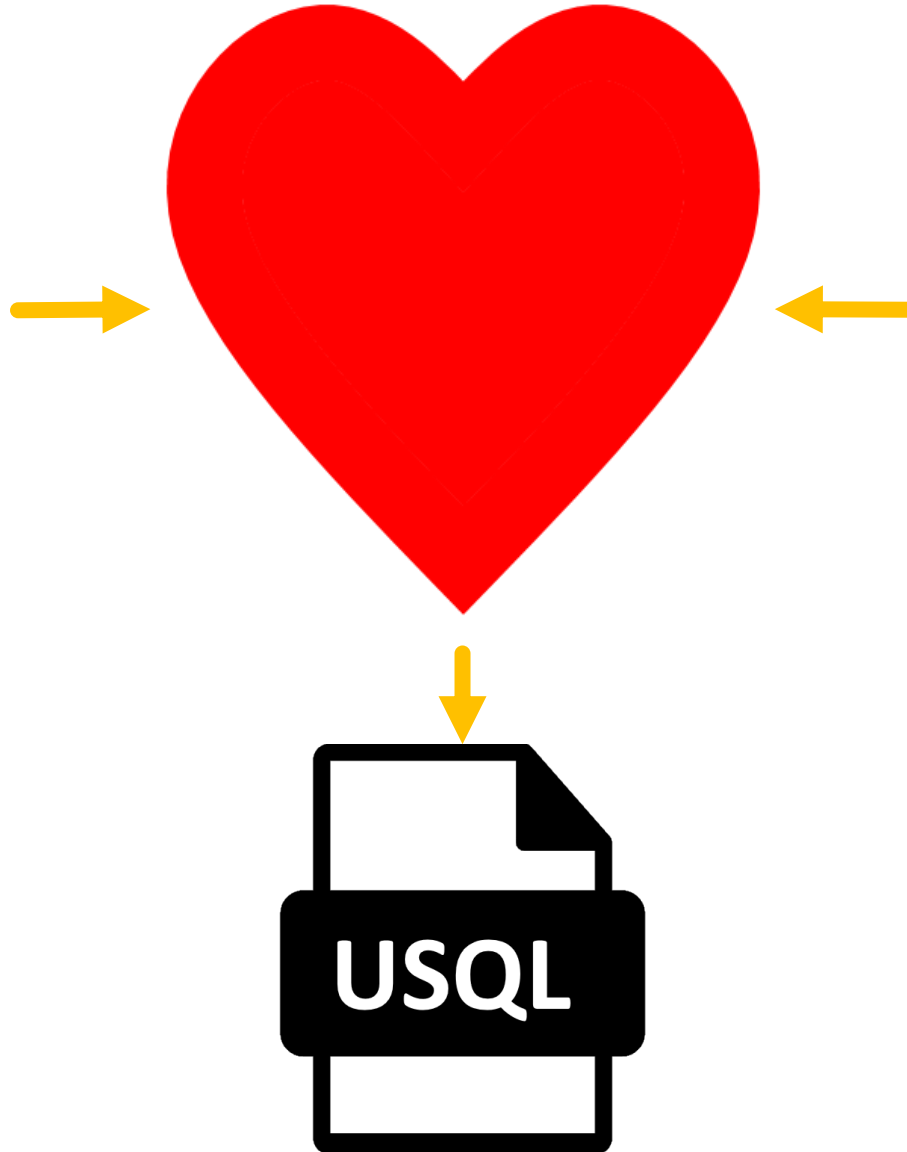


What is U-SQL?



What is U-SQL?

```
SELECT
    Domain,
    COUNT(*) AS Qty
FROM
    @Domains
GROUP BY
    Domain;
```



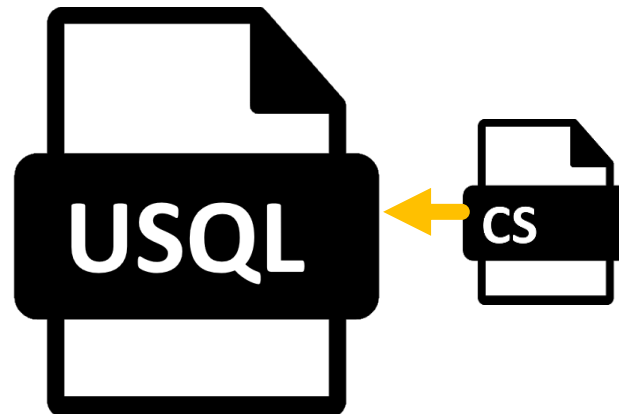
```
using System;

namespace USQLSampleApplication
{
    0 references | 0 changes | 0 authors, 0 changes
    public class CustomMethods
    {
        0 references | 0 changes | 0 authors, 0 changes
        static public int testMethod()
        {
            string testString = String.Empty;
            int testInt = Int32.MinValue;

            return testInt;
        }
    }
}
```

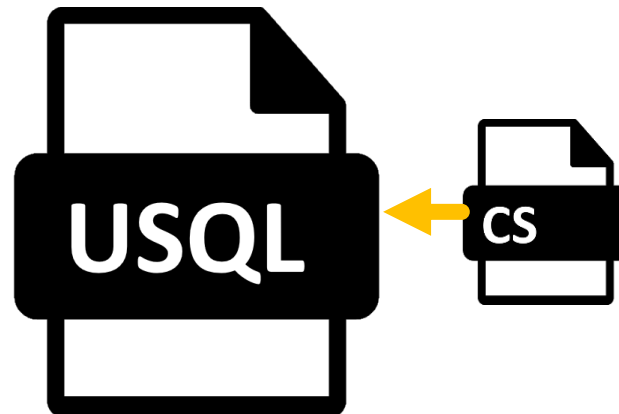

SELECT

```
@SizeAndCount =  
    SELECT  
        [ModifiedDate].ToString("yyyy") AS Year,  
        [FileName].Substring([FileName].IndexOf(".") + 1, 3) AS FileExtension,  
        COUNT(0) AS RecordCount,  
        Math.Ceiling(Convert.ToDecimal(SUM([Size]))) AS FileSizeTotalsMB,  
        Math.Ceiling(Convert.ToDecimal(SUM([Size])/1024)) AS FileSizeTotalsGB  
    FROM  
        @Raw  
    WHERE  
        [ActualFileName] == "FileDetailsTest.csv"  
    GROUP BY  
        [ModifiedDate].ToString("yyyy"),  
        [FileName].Substring([FileName].IndexOf(".") + 1, 3);
```



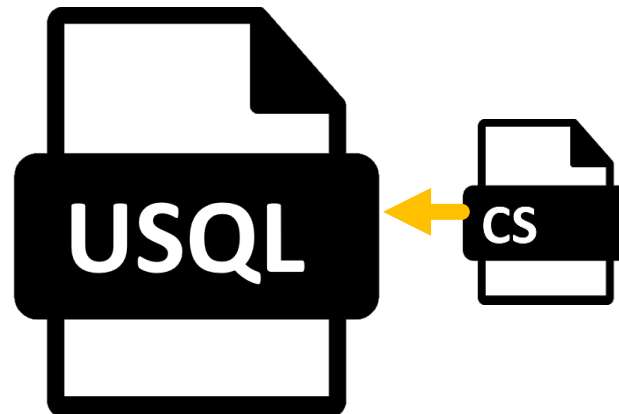
EXTRACT

```
@Base =  
    EXTRACT  
        [Field1] string,  
        [Field2] int  
    FROM  
        "BASE/External/AdventureWorks"  
    USING  
        Extractors.Csv();
```



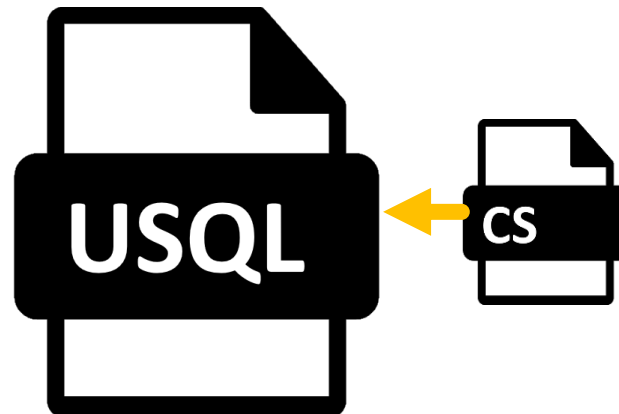
EXTRACT

```
@Base =  
    EXTRACT  
        [Field1] string,  
        [Field2] int  
    FROM  
        "BASE/External/AdventureWorks"  
    USING  
        Extractors.Csv(skipFirstNRows : 1, quoting : true,  
                        nullEscape:"\\N", encoding : Encoding.ASCII);
```



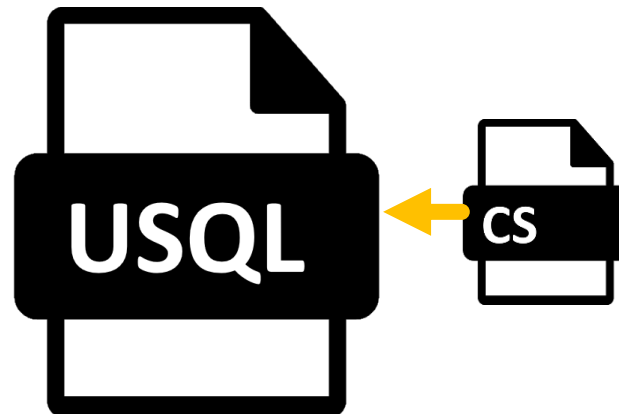
EXTRACT

```
@Base =  
    EXTRACT  
        [Field1] string,  
        [Field2] int,  
        //virtual columns  
        [FileName] string  
    FROM  
        "BASE/External/AdventureWorks/{FileName}"  
    USING  
        Extractors.Csv(skipFirstNRows : 1, quoting : true,  
                        nullEscape:"\\N", encoding : Encoding.ASCII);
```



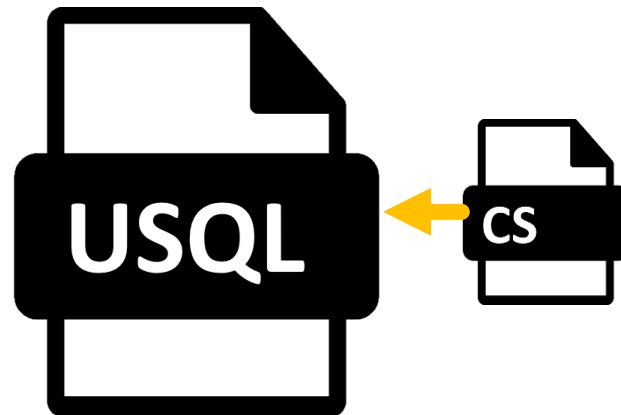
EXTRACT

```
DECLARE @Path string = "BASE/External/AdventureWorks/{FileName}";
@Base =
    EXTRACT
        [Field1] string,
        [Field2] int,
        //virtual columns
        [FileName] string
    FROM
        @Path
    USING
        Extractors.Csv(skipFirstNRows : 1, quoting : true,
            nullEscape:"\\N", encoding : Encoding.ASCII);
```



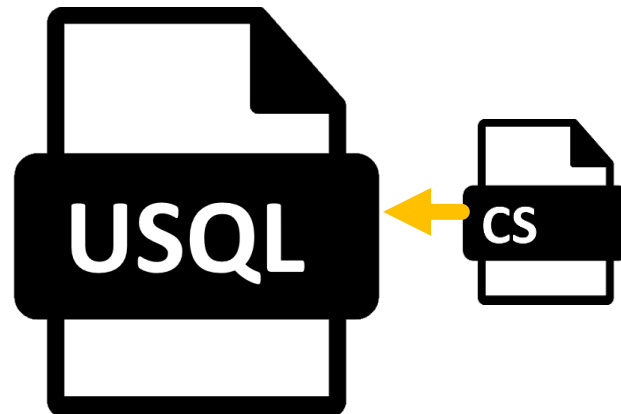
OUTPUT

```
OUTPUT @CompleteData  
TO "ENRICHED"  
USING Outputter.Csv();
```



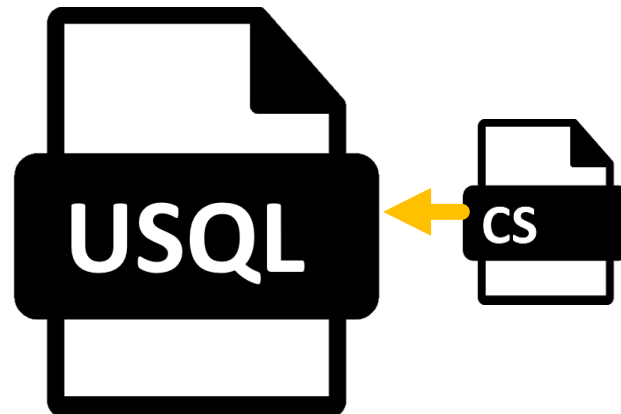
OUTPUT

```
OUTPUT @CompleteData  
TO "ENRICHED"  
USING Outputter.Csv(outputHeader : true);
```

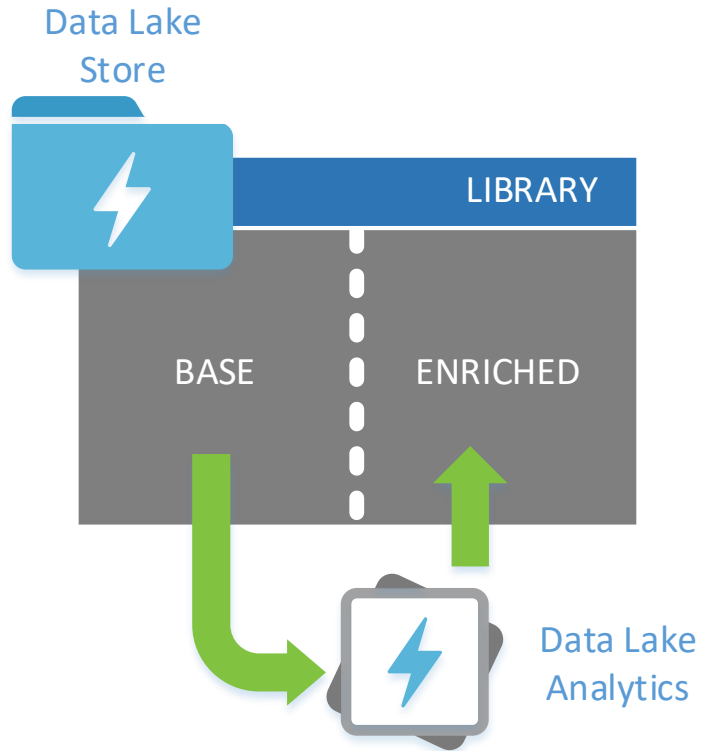


OUTPUT

```
DECLARE @Path string = "ENRICHED";  
OUTPUT @CompleteData  
TO @Path  
USING Outputter.Csv(outputHeader : true);
```



EXTRACT > SELECT > OUTPUT



```
@Base =  
    EXTRACT  
        [Field1] string,  
        [Field2] int  
    FROM  
        "BASE/External/AdventureWorks"  
    USING  
        Extractors.Csv();  
  
@CompleteData =  
    SELECT * FROM @Base;  
  
OUTPUT @CompleteData  
    TO "ENRICHED"  
    USING Outputter.Csv();
```

Module Agenda

What is U-SQL

Extract
Select
Output

Logical Instance

Table
Views
Stored Procedures

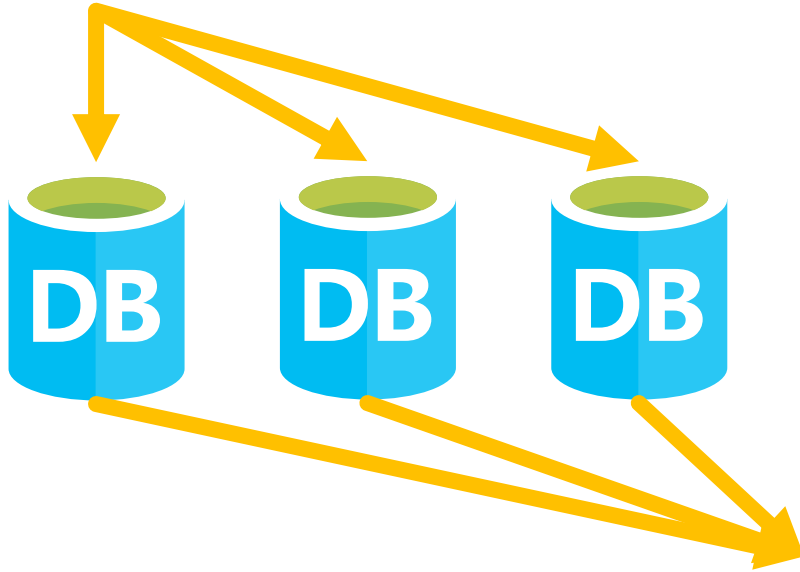
Data Lake Analytics As A Logical Instance



catalog



CREATE DATABASE IF NOT EXISTS BaseOfData01



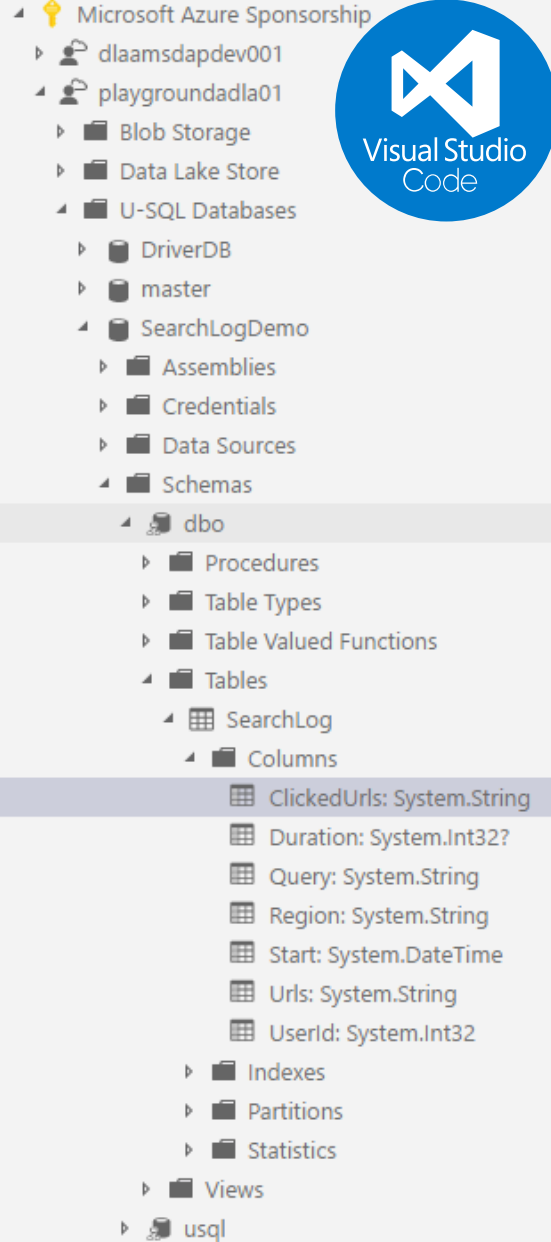
```
@AllData =  
    SELECT * FROM [Database1].[Schema1].[Table1]  
    UNION ALL  
    SELECT * FROM [Database2].[Schema2].[Table2]  
    UNION ALL  
    SELECT * FROM [Database3].[Schema3].[Table3];  
  
OUTPUT @AllData TO "BASE" USING Outputters.Csv();
```

Data Lake Analytics As A Logical Instance



catalog

```
CREATE DATABASE IF NOT EXISTS BaseOfData01
CREATE ASSEMBLY IF NOT EXISTS ImageCommon
CREATE PROCEDURE IF NOT EXISTS StoredProc01
CREATE SCHEMA IF NOT EXISTS Schema01
CREATE TABLE IF NOT EXISTS Table01
CREATE VIEW IF NOT EXISTS View01
CREATE CLUSTERED INDEX Index01 ON Table01
CREATE STATISTICS Stats01 ON Table01
```





U-SQL Further Reading

Microsoft U-SQL Language Reference Guide

<https://msdn.microsoft.com/en-us/azure/data-lake-analytics/u-sql/u-sql-language-reference>

SQL Server Central Stairway (21 chapters)

<http://www.sqlservercentral.com/stairway/142480/>

Stack Overflow U-SQL Tag

<http://stackoverflow.com/questions/tagged/u-sql>

MrPaulAndrew.com

<https://mrpaulandrew.com/tag/u-sql/>

Adatis Blogs

<http://blogs.adatis.co.uk/search?q=U-SQL>

Agenda for the Day

Module 1

Microsoft Azure

Module 2

Storage
Uploading Data
Data Lake

Module 3

Real-time Data
Streaming
Power BI

Module 4

U-SQL - Data
Transformation
Basics

Module 5

USQL - Advanced
Analytics
Cognitive Services

Module 6

Data Factory
Orchestration
Dynamic Pipelines

Module 7

Data Presentation
& Consumption
Power BI Models

Module 8

Other Services
Q&A