



2018

Global **Azure**
BOOTCAMP

Go Big or Go

- Home: Big Data Analytics & Azure Data Lake



Niels Berglund

Software Architect - Derivco

Microsoft Data Platform MVP

niels.it.berglund@gmail.com

www.nielsberglund.com

[@nielsberglund](#)

What is a Data Lake?

“A single store of all data... ranging from raw data (which implies exact copy of source system data) to transformed data which is used for various forms including reporting, visualization, analytics, and machine learning”

Azure Data Lake

- Comprehensive, big-data storage + analytics platform
- Purpose-built from real-world experiences
 - Office 365, Skype, Bing, etc.
- Leverage existing skills and technologies
- Benefits of an Azure-hosted service
 - Elastic, dynamically provisioned compute resources for varying query needs
 - Infinite storage capacity
 - Focus on extracting meaning from data, not on infrastructure

Azure Data Lake Store

- HDFS-as-a-service
- Durable, redundant storage
- A variety of data scenarios
 - High capacity
 - High frequency
 - High throughput
- Store data in its native format
 - Structured, semi-structured, unstructured storage formats



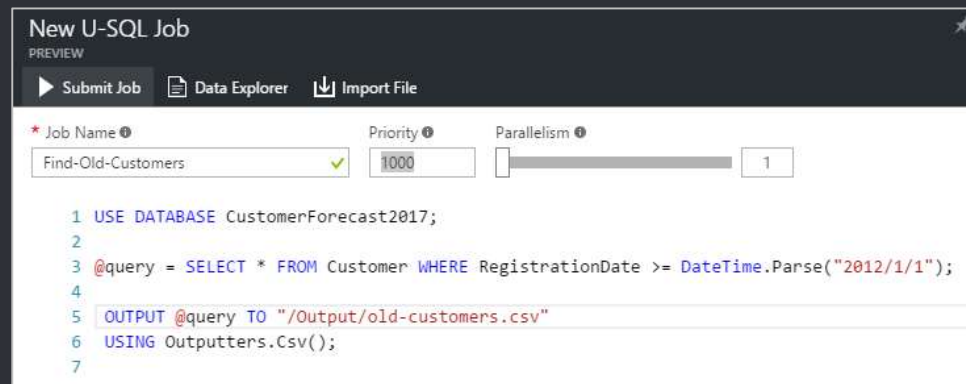
- Demo/Lab

Azure Data Lake Analytics

- Low-barrier alternative (or complement) to HDInsight and Hadoop ecosystem
- Scales dynamically to match data size and query complexity
- Built on Apache YARN
- Unit of interaction is an analytics job
 - Elastic infrastructure management is abstracted away
- U-SQL: query language rooted in SQL and C#

U-SQL

- Based on SQL and C#
 - C# expressions and types
 - Tables, views, window functions, etc.
 - User-defined functions/operators/aggregators in C#
- Typical job
 1. Read data from file/table/federated source
 2. Transform rowset in a pipeline
 3. Output rowset to table or file



```
1 USE DATABASE CustomerForecast2017;
2
3 @query = SELECT * FROM Customer WHERE RegistrationDate >= DateTime.Parse("2012/1/1");
4
5 OUTPUT @query TO "/Output/old-customers.csv"
6 USING Outputters.Csv();
7
```

- Demo/Lab

Thanks!

ANY QUESTIONS?

You can find me at

niels.it.berglund@gmail.com

www.nielsberglund.com

[@nielsberglund](#)