Summary

Robert C. Cain, MVP

OWNER, ARCANE TRAINING AND CONSULTING

@arcanecode www.arcanecode.com

Agenda

- Introduction
- 80% of the Operators You'll Ever Use
- Scalar Operators
- Advanced Aggregations
- Working with Datasets
- Time Series
- Machine Learning
- Exporting Data
- Summary

Data Types

Type	Other Names	.NET Type	gettype()	Kusto Internal Type
bool	boolean	System.Boolean	Int8	I8
datetime	date	System.DateTime	datetime	DateTime
dynamic		System.Object	array or dictionary	Dynamic
guid	uuid, uniqueid	System.Guid	guid	UniqueID
int		System.Int32	int	I32
long		System.Int64	long	I64
real	double	System.Double	real	R64
string		System.String	string	stringbuffer
timespan	time	System.TimeSpan	timespan	TimeSpan

As of May 2018 GUID implementation is not complete. Suggest using string instead
http://bit.ly/kqldatatypes

Best Practices – Do's

- Use time filters first
- Next, use filters that remove the majority of the data
- Use filters as early as possible, before using extend
- Prefer *has* over *contains*, as has is more performant
- Look in specific columns rather than using wildcards (*)
- When using join, make the table with fewer rows come first (the left table)
- When using join, project only needed columns from both sides
- When extracting fields from dynamic objects across millions of rows, use materialize to reduce the impact of column extraction

Best Practices – Don’ts

- Do not run queries for the first time without limiting the results using a limit or count at the end. Otherwise you could return GB of data
- If you are applying conversions (JSON, strings, etc) to over 1 billion records, reshape your query to reduce the data fed into the conversion
- Don’t use tolower(Col) == “lowercasestring” to do case insensitive comparisons. Instead use Col =~ “lowercasestring”
- Don’t filter on a calculated column, if you can filter on a table column

What's Next

- More and more Microsoft services will be exposing their data via KQL
- Ability to integrate with external data sources
- Even more Machine Language capabilities
- Performance optimizations
- Continued expansion of the operator set

Summary

- Introduction
- 80% of the Operators You'll Ever Use
- Scalar Operators
- Advanced Aggregations
- Working with Datasets
- Time Series
- Machine Learning
- Exporting Data
- Summary