alteryx

Alteryx Analytics 11.0 Performance Benchmarking Results

v3, February 2017

Contents

Introduction

Alteryx Analytics is the leading platform for self-service data analytics. It provides analysts with the ability to easily prep, blend, and analyze all of their data using a repeatable workflow, then deploy and share analytics at scale for deeper insights.

To measure and monitor the processing speeds of Alteryx analytic capabilities, we conduct performance benchmark testing on an ongoing basis. Our benchmarking tests cover a wide variety of data preparation, data blending, and spatial analytics processes using different hardware configurations and datasets. Results can be used to measure and monitor system performance, identify problems, and help determine system deployment needs.

The goal of this report is to share our benchmarking performance test results for Alteryx Designer in order to inform you of Alteryx performance and help you determine your hardware needs for deploying Alteryx.

Summary of Results

Results show Alteryx performed a variety of data preparation and analytics processes within seconds or minutes depending on the complexity of analysis and number of records processed.

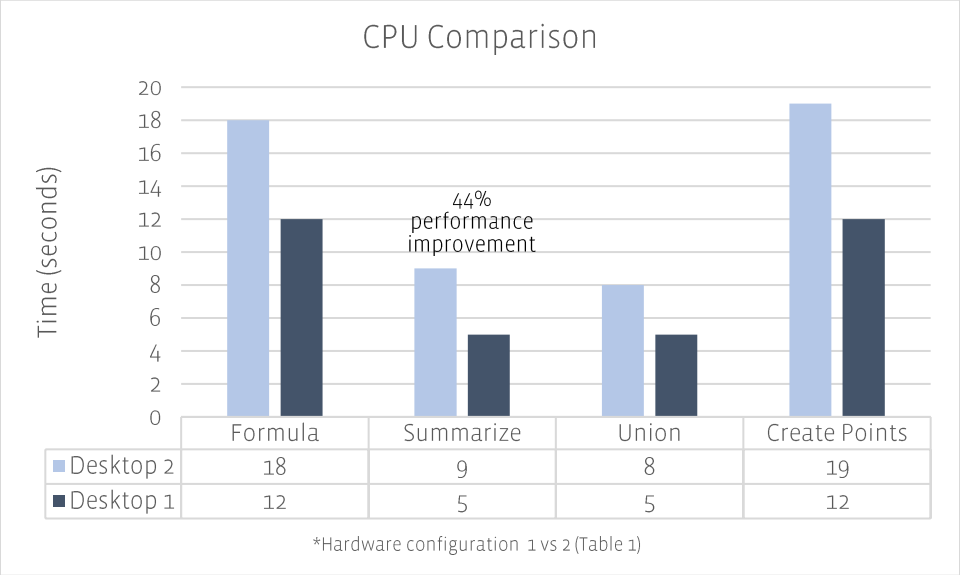
Benchmark tests were created as Alteryx workflows and executed using Alteryx Designer 11.0 with two generic datasets of different sizes. We ran each test multiple times on different hardware configurations.

Our tests revealed three key results.

Result #1: Improved Performance with Faster CPU

Alteryx performance improved when run on a machine with a faster central processing unit (CPU).

We tested Alteryx performance on two machines with processing speeds of 3.6 GHz and 2.5 GHz. When using the machine with a processing speed of 3.6 GHz, Alteryx performance improved up to 44%. Performance improvements varied based on the tools used in each workflow and the number of records processed. The chart shows a sample of the results.

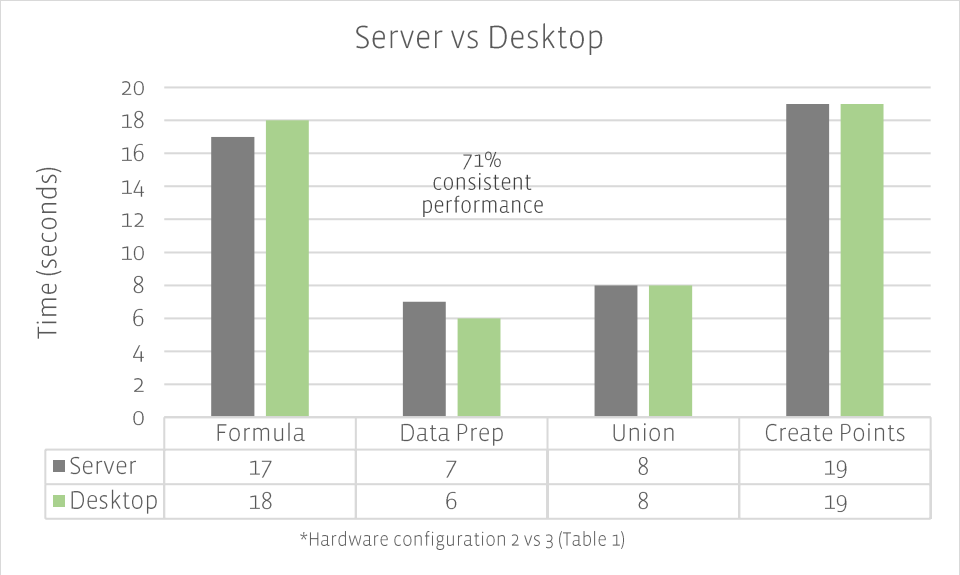


Result 1

Result #2: Similar Performance on Server & Desktop

Alteryx Designer installed on a server performed as fast as Alteryx Designer installed on a desktop.

We compared the performance of Alteryx Designer installed on a server and a desktop with similar hardware specifications (see Table 1). In 71% of tests, Alteryx Designer installed on a server and a desktop performed tests at the same speed or with less than a 5 second difference. The chart shows a sample of the results.



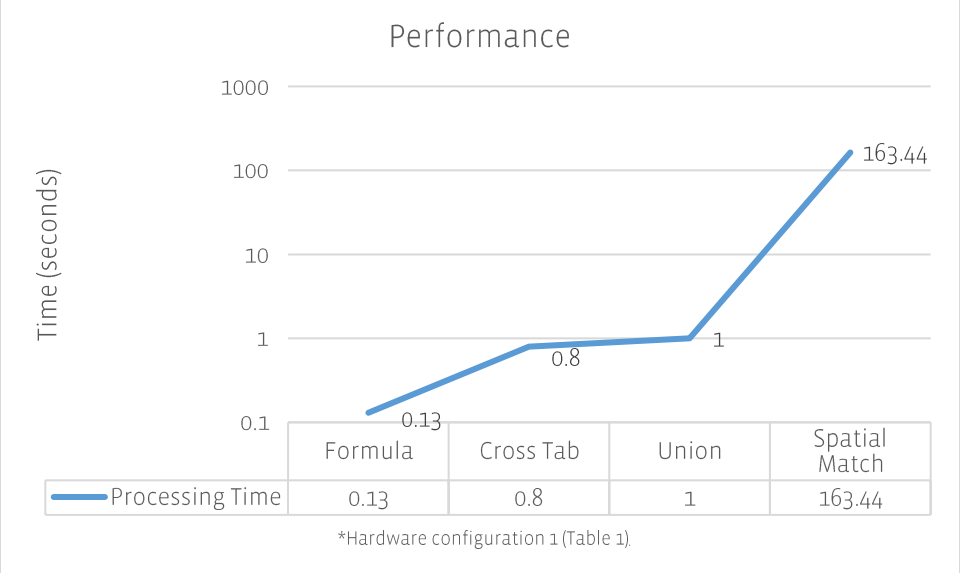
Result 2

Result #3: Increased Processing Times with Complexity of Analysis

Alteryx processing times increased with the complexity of analysis performed.

We tested the time it took Alteryx to prepare, transform, and blend data as well as to perform spatial analytics processes. Tests included one tool from each of those groups. For each test, we used the results to calculate the approximate time it would take Alteryx to process 1 million records.

Overall, complex processes such as blending data and performing spatial analytics required more time than data preparation tasks such as adding a column to a dataset or sorting data in a dataset. The chart shows a sample of the results.



Result 3

Benchmarking Approach

System Hardware Specifications

Test were conducted on three different hardware configurations. Alteryx Designer 11.0 was installed on each configuration. Results were compared to measure the impact of hardware configurations on Alteryx performance.

Configuration	1 Desktop Virtual Machine	2 Desktop Virtual Machine	3 Server Virtual Machine
Processor	Intel® Core™ i7-4790 Dual Processor	Intel® Xeon® E5-2680 v3 Dual Processor	Intel® Xeon® E5-2680 v3 Dual Processor
CPU	3.6 GHz	2.5 GHz	2.5 GHz
Memory	16 GB	32 GB	32 GB
Disk size	450 GB	530 GB	600 GB
OS	Windows 7 Enterprise 64-bit	Windows 7 Enterprise 64-bit	Windows Server 2012 R2 Standard

Table 1

Tests and Data

Tests were created as Alteryx workflows designed to measure the time it takes Alteryx to perform data preparation, data transformation, data blending, and spatial analytics processes. Each test included at least one Input Data tool and a specific tool or set of tools from one of the test groups.

Tests were also designed to measure the time it takes Alteryx to process datasets of different sizes. We ran each test using generic datasets of two sizes saved as .yxdb files. Additionally, for each test, we used the results to calculate the approximate time it would take Alteryx to process 1 million records.

Data Processing Tests

Tests in this group examined the performance of tools used to prepare, transform, and blend data.

Test #	Test Name	# Records Set 1	# Records Set 2
1	Formula	90,000,000	200,000,000
2	Filter	5,000,000	200,000,000
3	Sort	5,000,000	200,000,000
4	Data Preparation	5,000,000	105,000,000
5	Summarize	90,000,000	150,000,000
6	Cross Tab	5,000,000	200,000,000
7	Union	5,000,000	200,000,000
8	Join	3,000,000	1,000,000,000

Table 2

Spatial Analytics Tests

Tests in this group examined the performance of tools that use functionality such as geocoding, point creation, trade area, and spatial matching.

Test #	Test Name	# Records Set 1	# Records Set 2
9	Create Points	90,000,000	150,000,000

10	Spatial Match	2,000,000	115,000,000
11	Spatial Pie Wedge	90,000,000	150,000,000

Table 3

Test Execution

Alteryx Designer 11.0 was installed on three virtual machines of different hardware specifications. Test were run one at a time using the Alteryx Engine command line: *AlteryxEngineCmd.exe [MyWorkflowName].yxmd*.

Each test was executed three times with two datasets of different sizes on each hardware configuration. Because Alteryx holds files and references in memory, subsequent runs of the same test may process faster than the initial test. Since Alteryx was not restarted between each test, we chose to average the test results.

Benchmarking Test Results

The following results show the average time it took Alteryx Designer 11.0 installed on hardware configuration 1 (see Table 1) to execute each test using datasets of two sizes. Results also include the calculated time it would take Alteryx to process 1 million results.

Data Processing Tests

Data Processing tests are divided into three groups: Data Preparation, Data Transformation, and Data Blending. These tests include one or more Alteryx tools used to perform data processing tasks including sorting, averaging, and joining data.

Data Preparation Tests

Tests in this group identified the time it took Alteryx to perform common data preparation tasks including filtering and sorting data, and adding new columns to a dataset. Results show Alteryx performed data preparation processes on 100 million records in one second or less.

Test #1: Formula

The Formula tool can be used to create or update fields using one or more expressions to perform a variety of calculations and operations. This test identified the time it took to apply a simple expression to add one field value to another in order to create a new field.

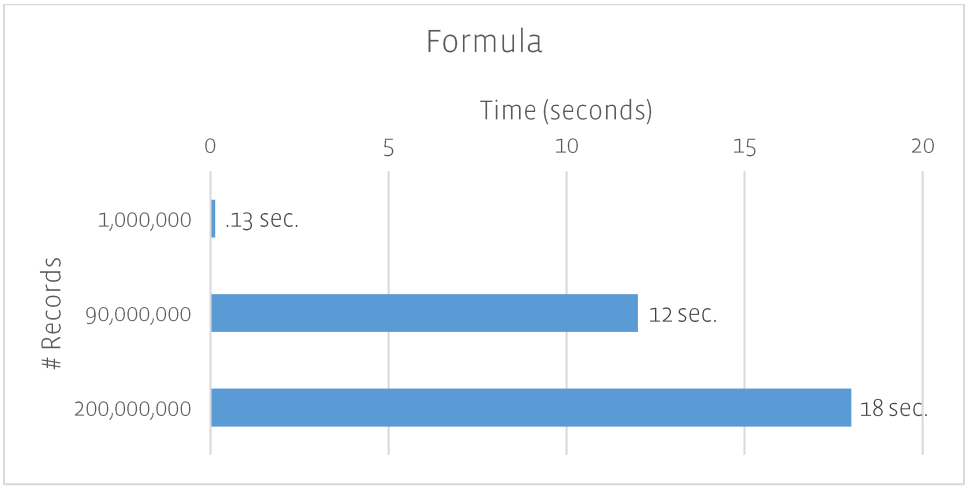


Table 5

Test #2: Filter

The Filter tool can be used to find records in a file that meet specified criteria. The Filter test identified the time it took to find all records in a dataset with a specific field equaling a specific value, using the Filter tool.

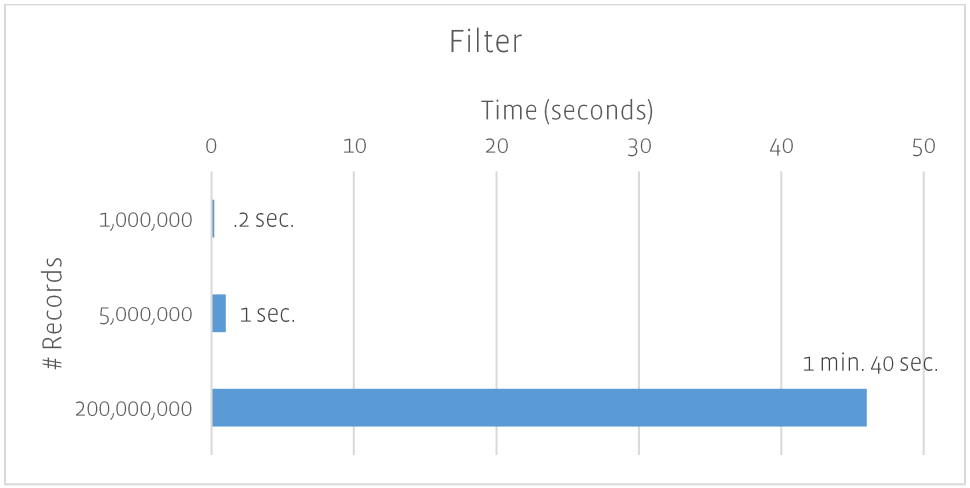


Table 6

Test #3: Sort

The Sort test identified the time it took to sort values, in ascending order.

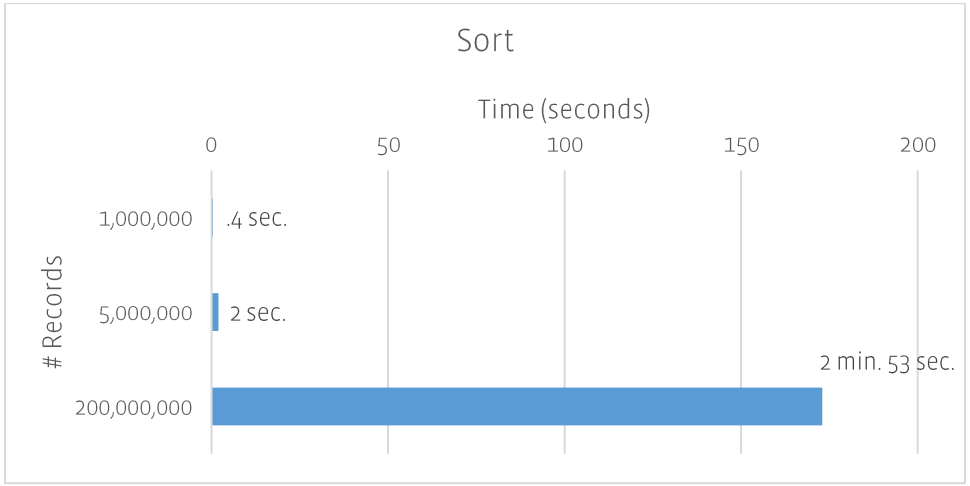


Table 7

Test #4: Data Preparation

This test identified the time it took to perform multiple data preparation tasks. The Data Preparation test contained multiple tools including: Filter, Multi-Formula, Cross Tab, Select, and Dynamic Rename. Tools were used to filter a dataset, select specific values within the set, cleanse data, and reformat then label results.

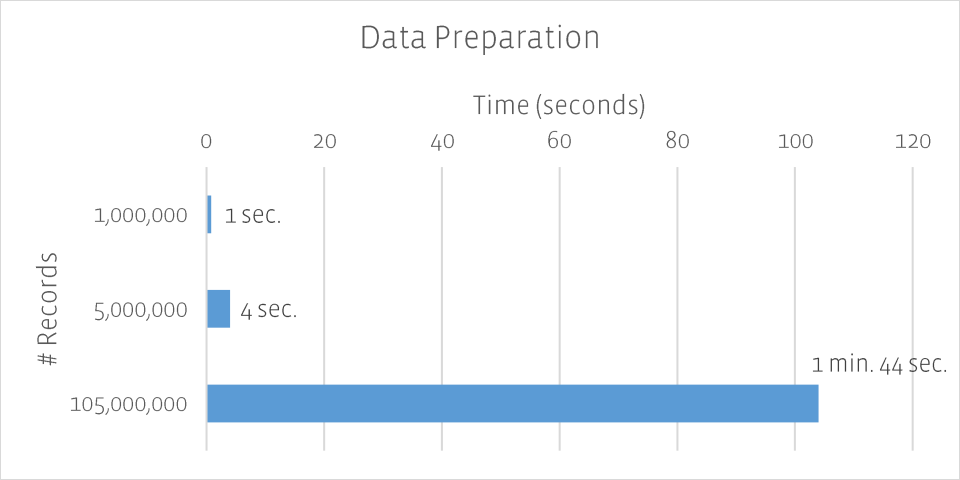


Table 8

Data Transformation Tests

Tests in this group identified the time it took Alteryx to average records in a dataset and change the orientation of data in a dataset. Results show Alteryx calculated the average of 150 million records in less than 10 seconds.

Test #5: Summarize

The Summarize tool can perform several summarize functions including group by, sum, count, string concatenation, and processing spatial data. This test identified the time it took to sum selected values and divide the sum by the total number of values using the Summarize tool.

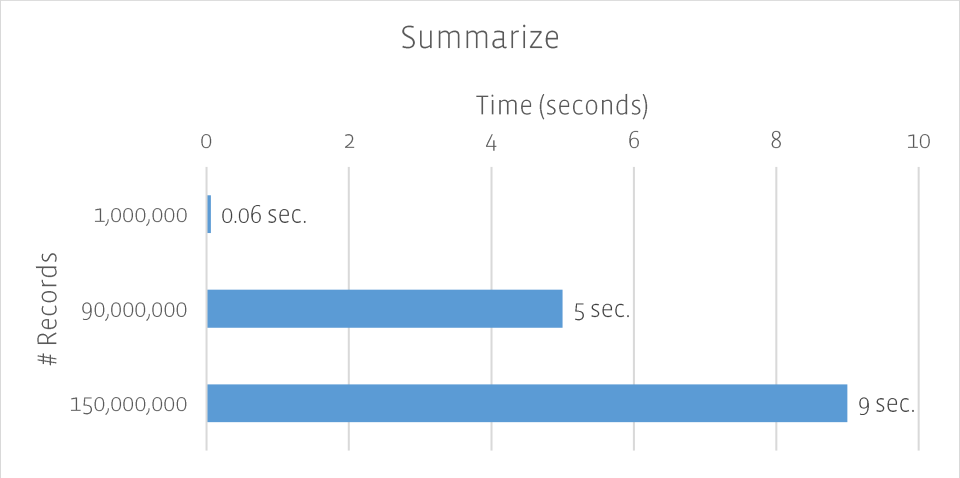


Table 9

Test #6: Cross Tab

The Cross Tab tool pivots the orientation of a dataset, switching the vertical and horizontal axes. This test identified the time it took to create new columns and sum the values within the columns.

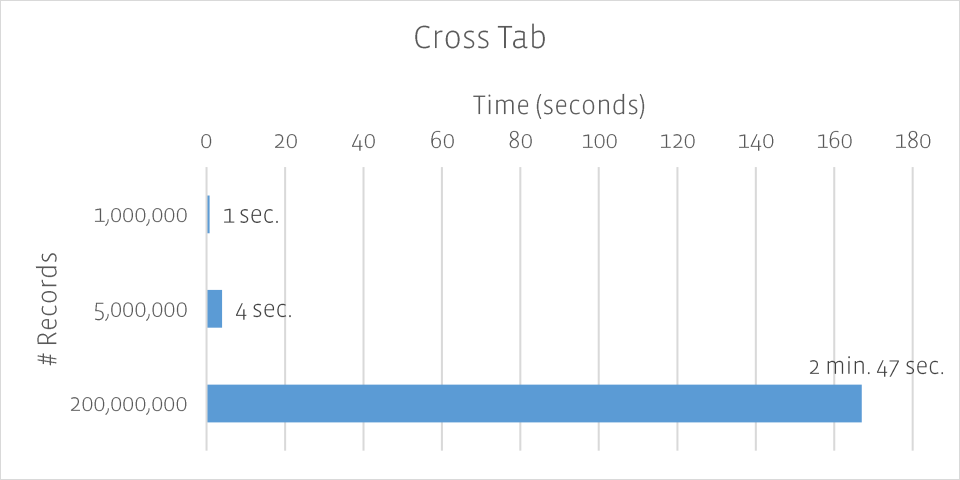


Table 10

Data Blending Tests

Tests in this group identified the time it took Alteryx to perform data blending processes. Data blending combines data from two datasets to create a single dataset. Results show Alteryx blended two datasets of 1 billion records in two minutes.

Test #7: Union

The Union tool combines datasets of similar structure into a unified set based on field names and positions. This test identified the time it took to combine two datasets based on a selected field name using the Union tool.

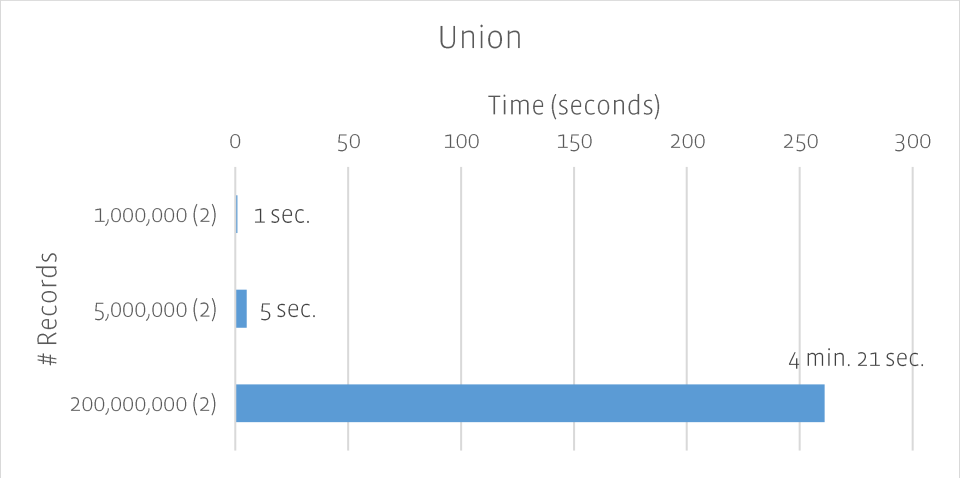


Table 11

Test #8: Join

The Join tool combines two datasets based on common fields or record positions. This test identified the time it took to join two datasets based on a common field.

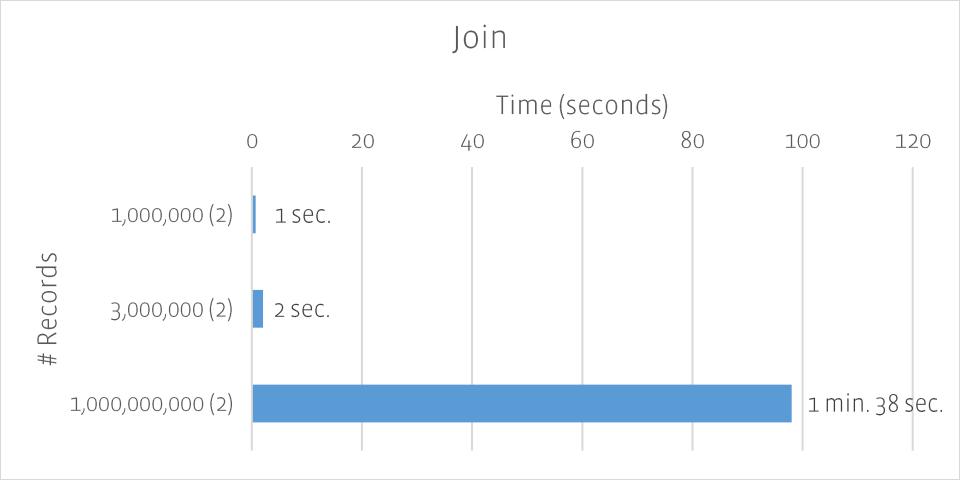


Table 12

Spatial Analytics Tests

Tests in this group identified the performance of spatial analytics tools used for geocoding, point creation, trade area, drive time, and spatial matching. Results show Alteryx performed spatial analytics processes on 1 million records in 2.5 minutes. Processing times increased with the number of records processed and the complexity of analysis performed.

Test #9: Create Points

The Create Points tool combines longitude and latitude to create a spatial object. This test identified the time it took Alteryx to create spatial points using numeric coordinate fields.

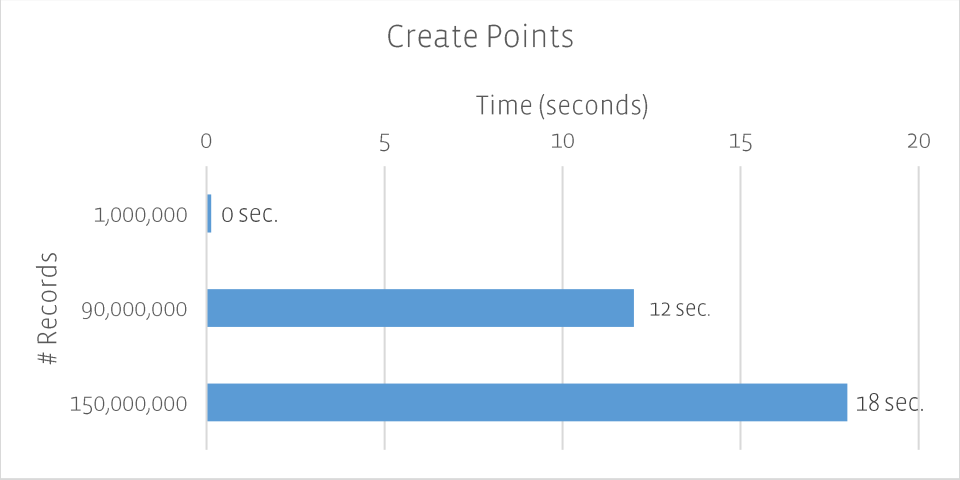


Table 13

Test #10: Spatial Match

This test identified the time it took to match records where two sets of spatial objects intersect and have an area in common, using the Spatial Match tool.

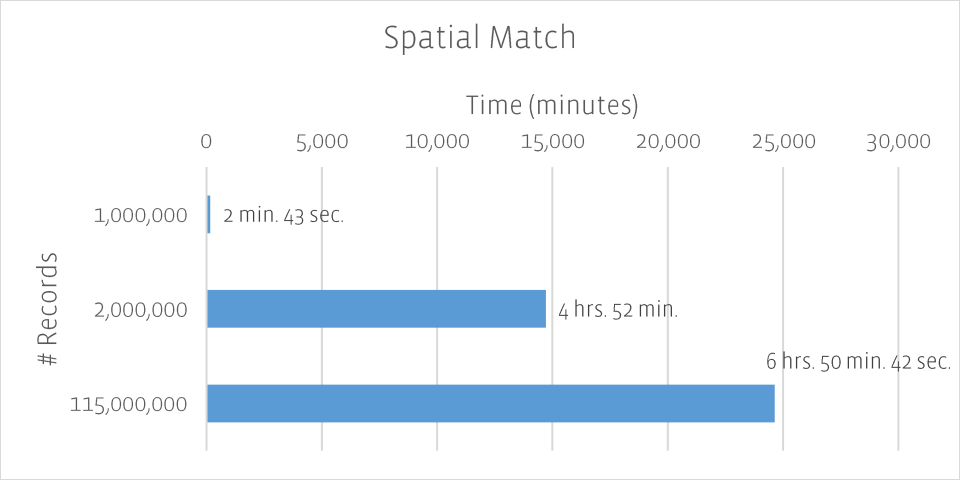


Table 14

Test #11: Spatial Pie Wedge

This test identified the time it takes Alteryx to perform multiple spatial analytics processes including creating a pie wedge visual using a dataset with area, width, and direction, from the center point of each wedge. The test contained several spatial analytics tools: Create Points, Trade Area, Poly-Split, Distance, and Poly-Build.

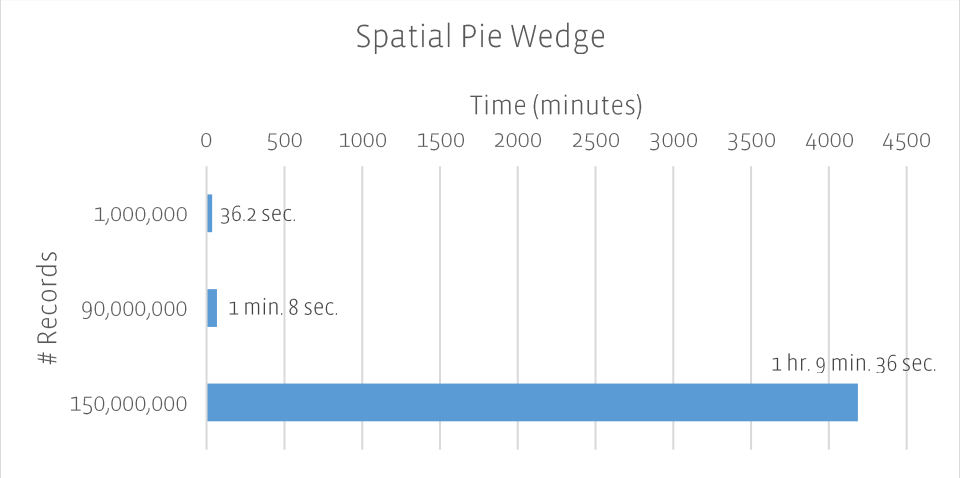


Table 15

All Tests

The following chart shows the calculated time to process 1 million records for each test.

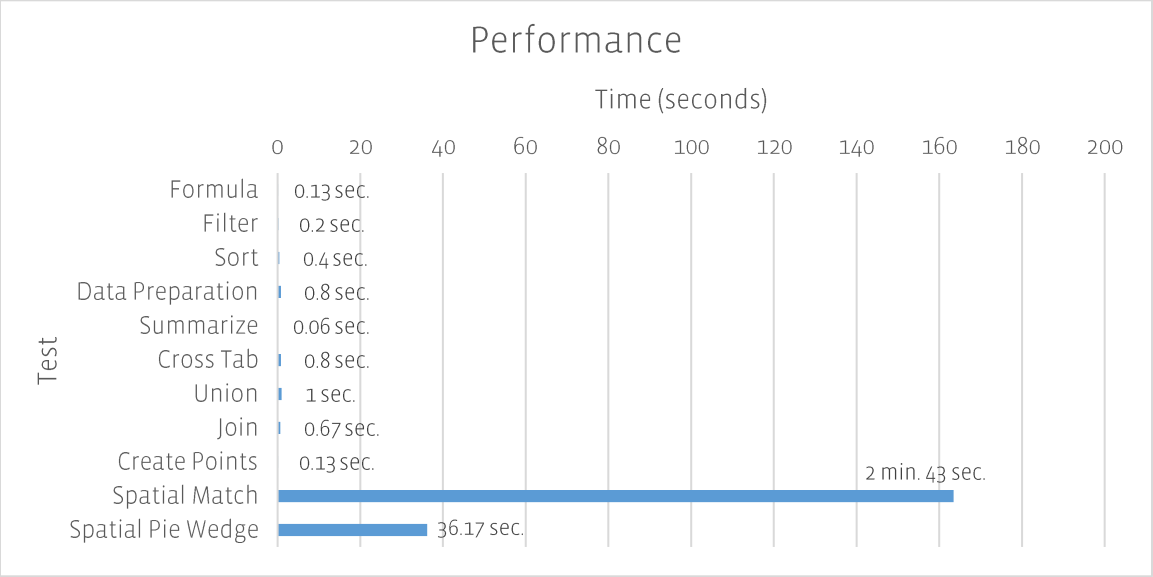


Table 17

Conclusion

These benchmark test results show Alteryx performed a variety of data preparation and analytics processes within seconds or minutes depending on the complexity of analysis and number of records processed. Alteryx performed processes at comparable speeds regardless of hardware configuration.

We conducted performance benchmark testing to measure and monitor the processing speed of Alteryx analytic capabilities in order to evaluate system performance, identify problems, and help determine system deployment needs. To accomplish this goal, we created tests that identified the time it took Alteryx Designer to perform data preparation, data blending, and spatial analytics processes using different hardware configurations and datasets.

Results are intended to inform you of Alteryx performance and help you determine your hardware needs for deploying Alteryx.