

A Data Analysis on Jobs Run by ACCRE

Data Science 5 Cohort
November 11, 2021



Agenda

- 1. Introduction**
2. Slurm Usage Trends
3. Slurm Job Outcomes
4. Slurm Partition Usage Breakdown
5. Slurm Unresponsiveness Analysis

Background

The Advanced Computing Center for Research and Education (ACCRE) requested an analysis on data on jobs run on ACCRE's hardware. Slurm scheduler is processing so many job completions frequently that it sometimes becomes unresponsive to commands. This is a particularly bad problem for clients who use automated submission systems, such as members of the Open Science Grid who submit jobs to slurm.

Goals

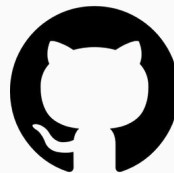
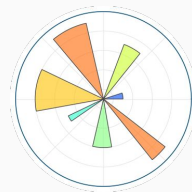
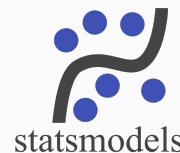
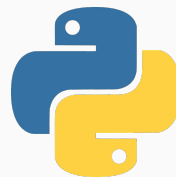
- Determine if high quantities of job completions in a short time period cause the scheduler to be unresponsive
- Determine the rough threshold at which it becomes an issue

Datasets and Tools



Datasets

- Job records of jobs processed by the Slurm scheduler
- Two server log files (CE5, CE6) of every Slurm command



What are we answering today?

- ★ How is Slurm utilized by users?
- ★ What are outcomes of jobs submitted to the Slurm scheduler?
- ★ How does Slurm usage compare by partition?
- ★ What causes the Slurm scheduler to become unresponsive?

Format

Agenda

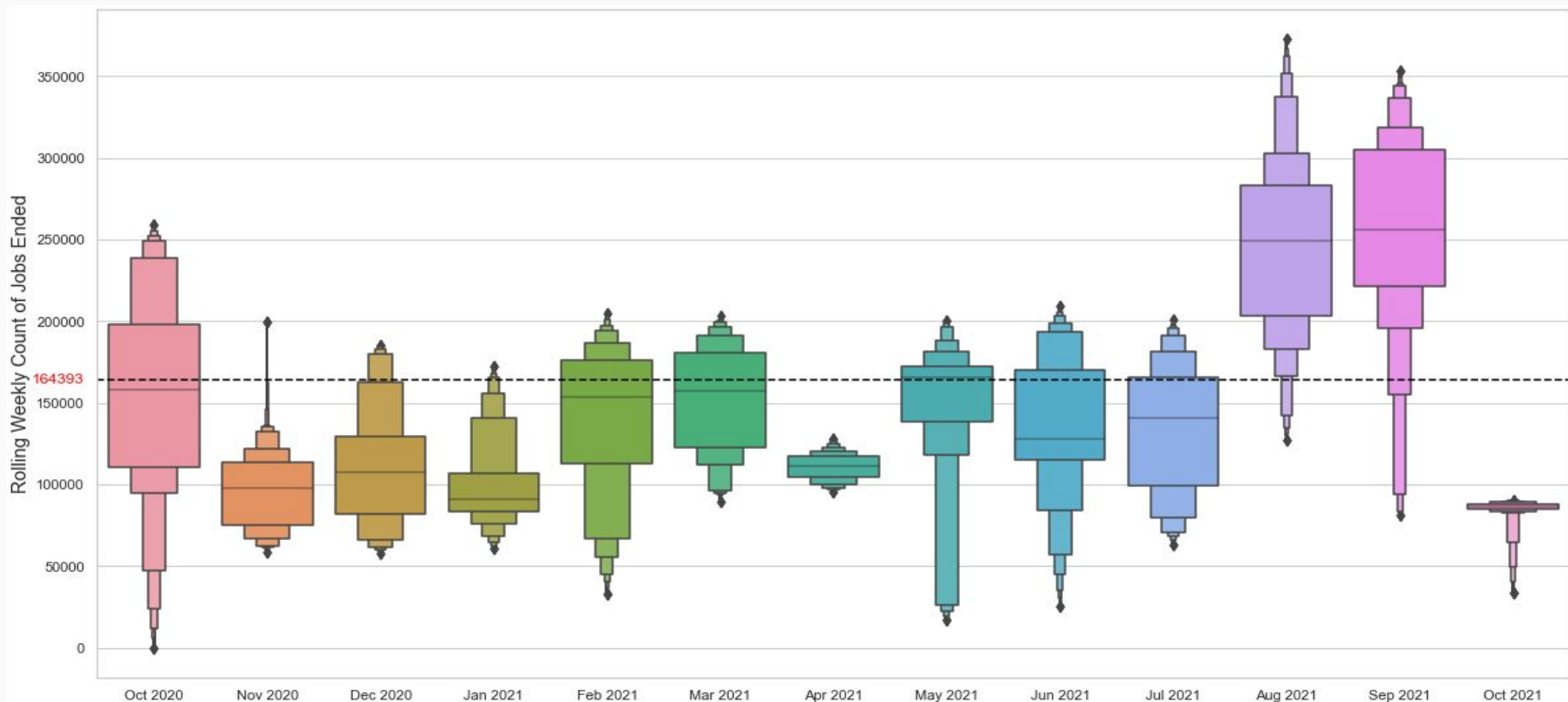
1. Background
2. **Slurm Usage Trends**
 - How many jobs does the Slurm scheduler process on a monthly/weekly/daily/hourly basis?
3. Slurm Job Outcomes
4. Slurm Partition Usage Breakdown
5. CE5/6 Server Log Analysis



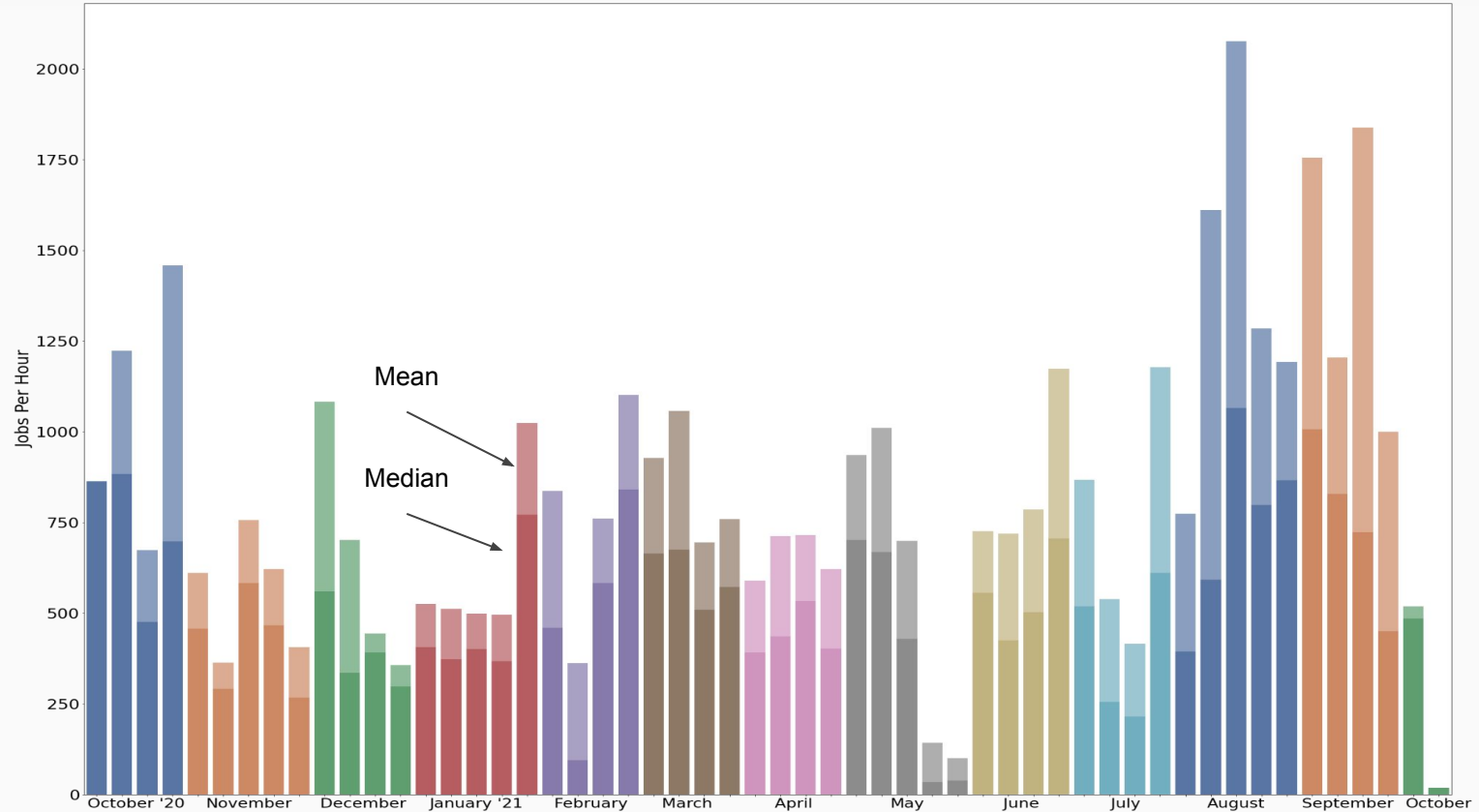
Agenda

1. Background
2. **Slurm Usage Trends**
 - How many jobs does the Slurm scheduler process on a monthly/weekly/daily/hourly basis?
3. Slurm Job Outcomes
4. Slurm Partition Usage Breakdown
5. Slurm Unresponsiveness Analysis

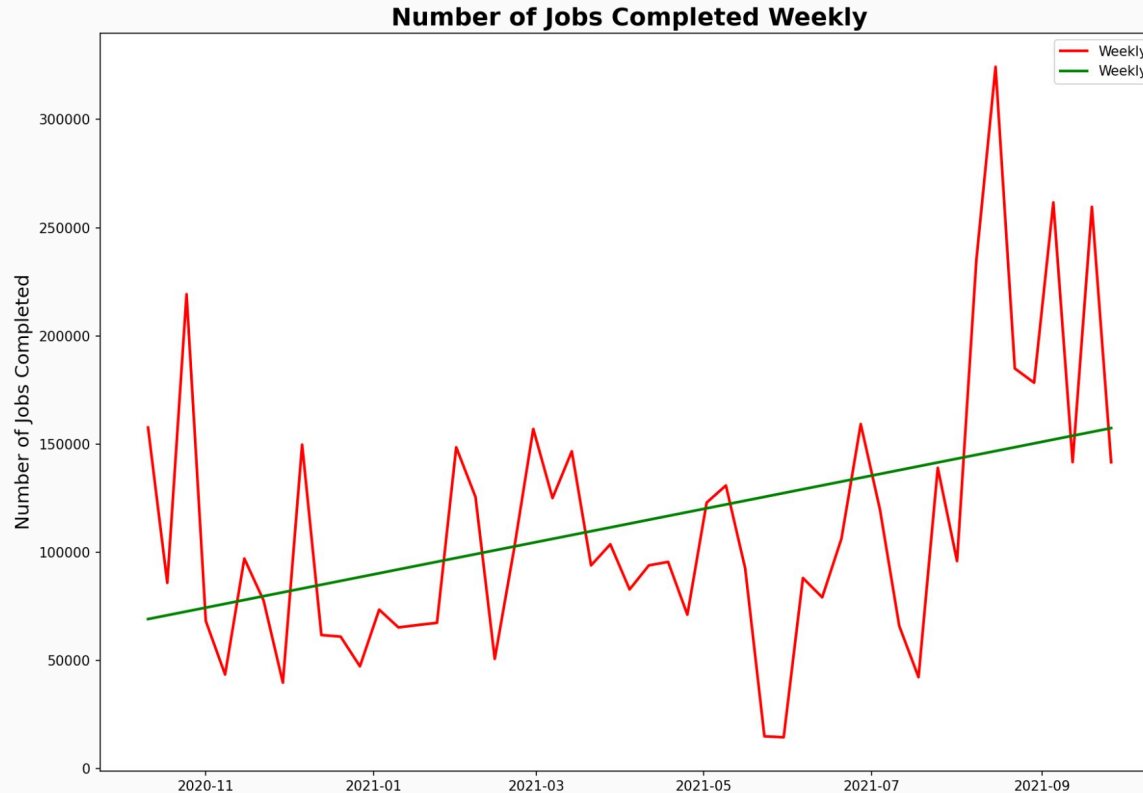
Monthly Trends: April is Strange, August and September Pick up the Pace



Weekly Trends: Certain hours each week see a higher number of jobs ended

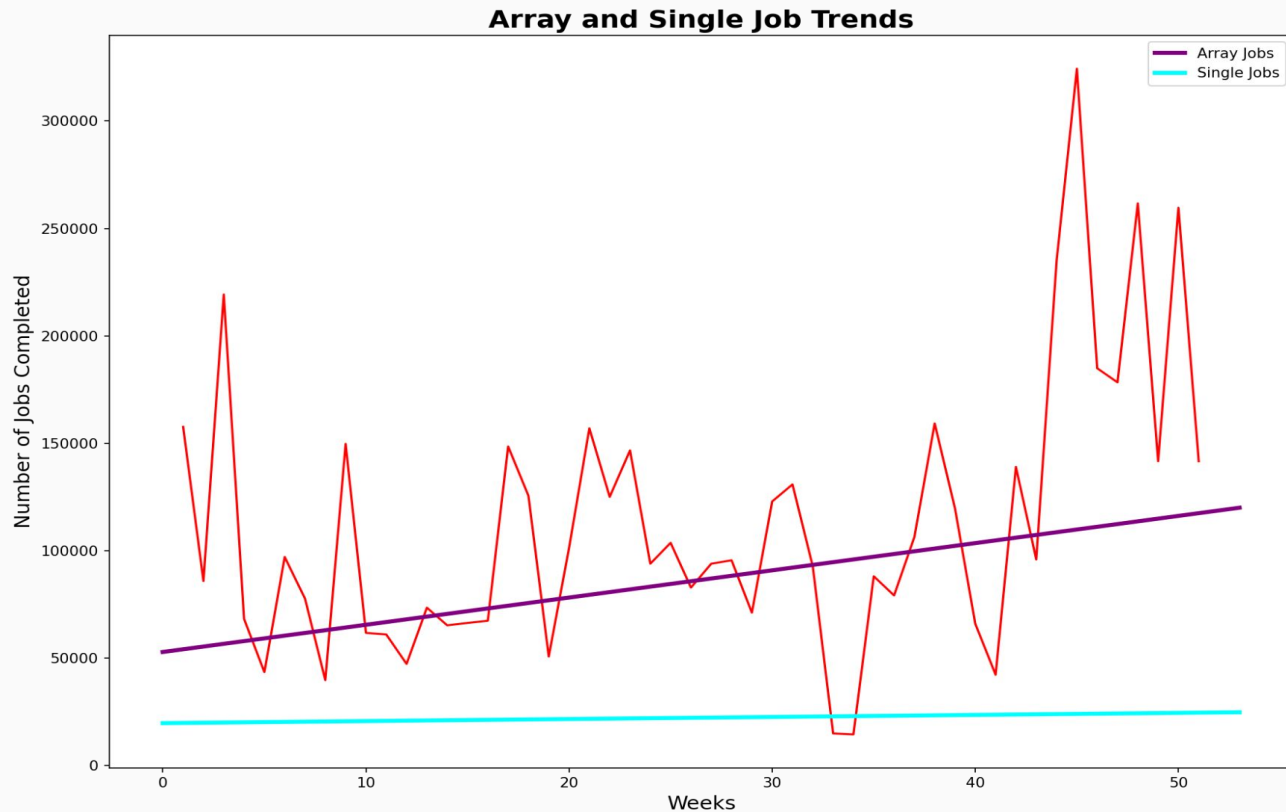


Weekly Trends: Completed Jobs per Week Trend Upward Over Time

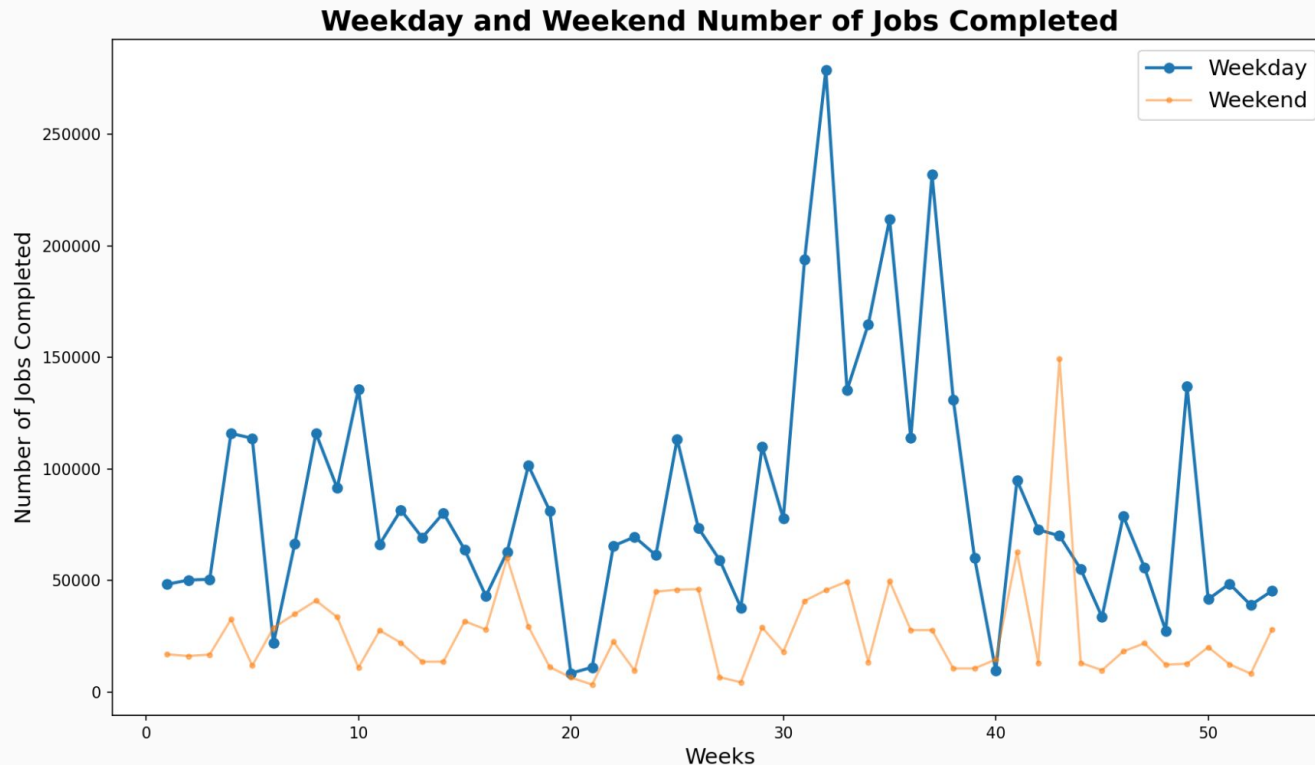


How many jobs does the Slurm scheduler process on a monthly/weekly/daily/hourly basis?

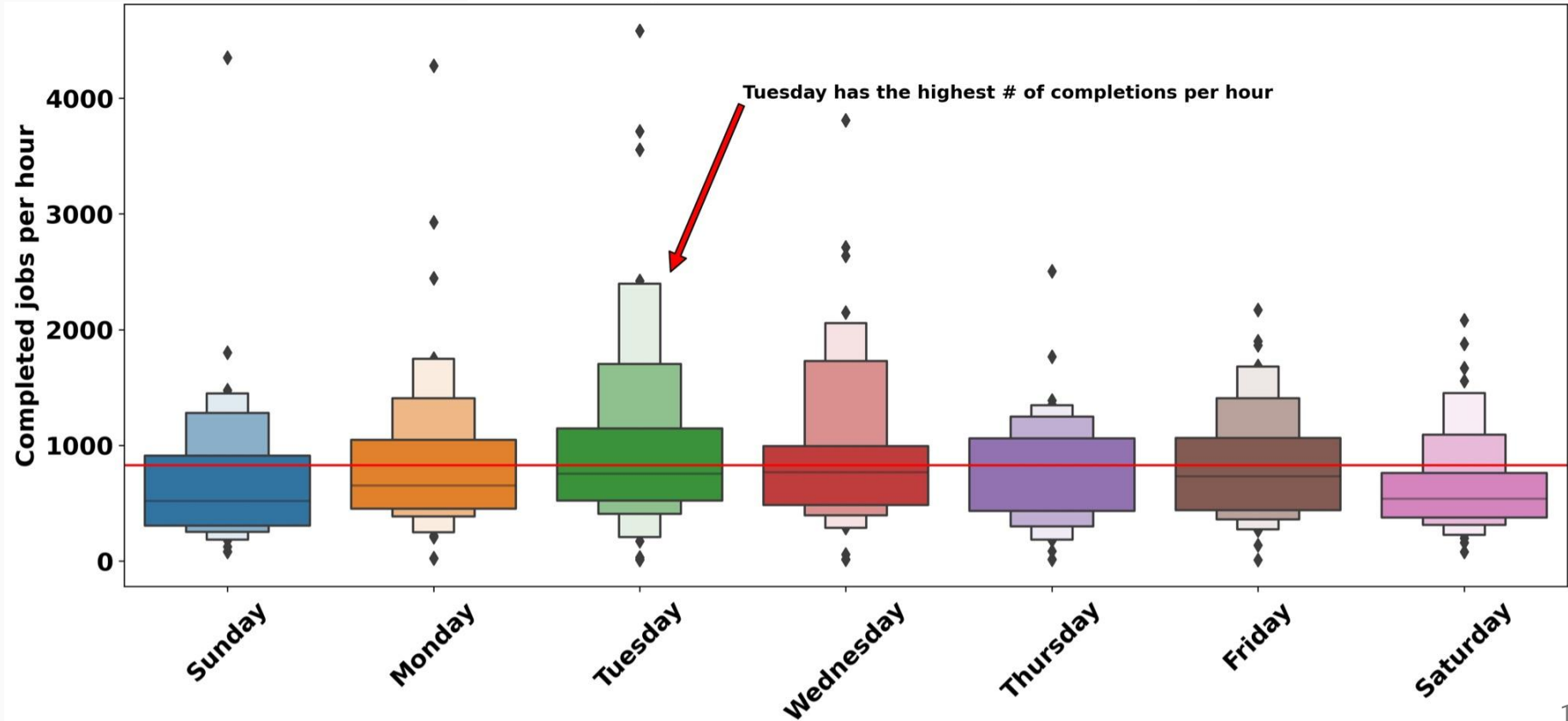
Weekly Trends: More Array Jobs Are Completed Over Time Than Single Jobs



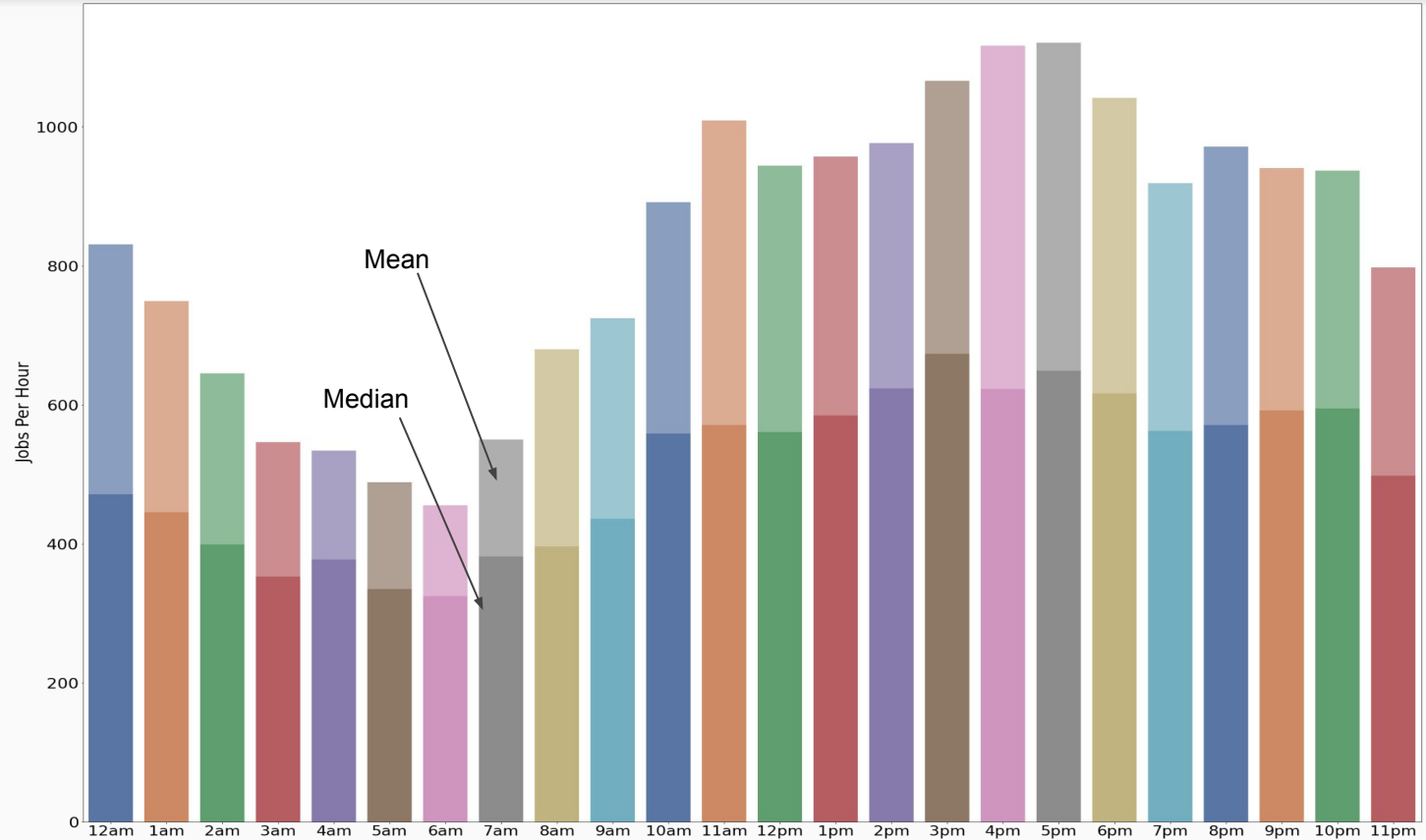
Weekly Trends: Job Completions are Higher on Weekdays than on Weekends



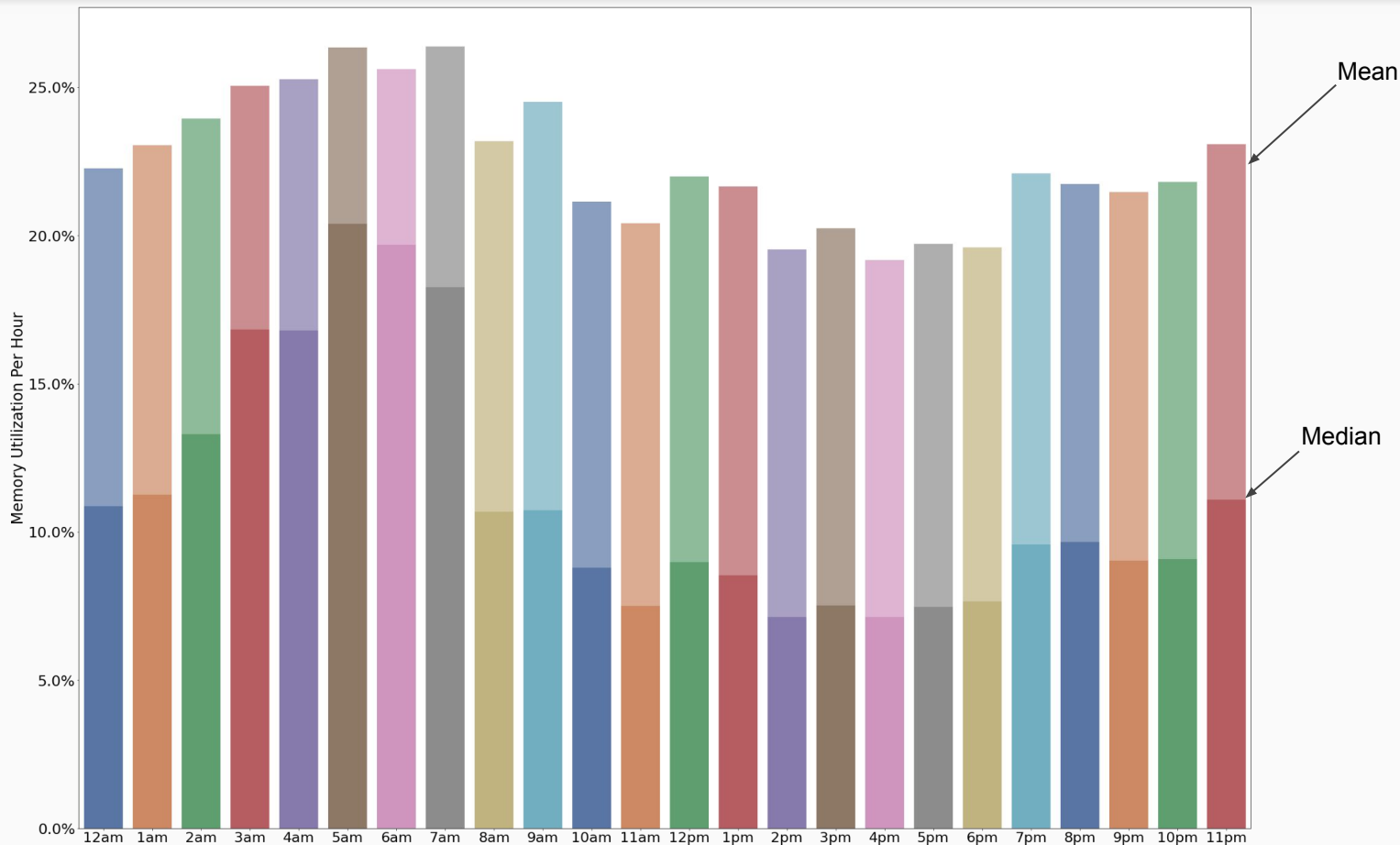
Daily Trends: Tuesday Has Highest Number of Jobs Completions per Hour



Hourly Trends: Higher Number of Jobs Ended from Midday to Evening Hours



Hourly Trends: Lower Memory Utilization Seen From Jobs Ended During Midday to Evening Hours



Hourly Trends: Job Lengths per Hour

Descriptive Statistics For Job Length Per Hour

	All Jobs Ended
Mean job length	5 hours 54 min
Median job length	26min
Min	-1 days 23 hours 2 min*
Max	34 days 23 hours 15 min

*69 jobs on November 1, 2020 have END times that occur before the BEGIN time. These jobs are in COMPLETED and FAILED states.

Agenda

1. Background

2. Slurm Usage Trends

- How many jobs does the Slurm scheduler process on a monthly/weekly/daily/hourly basis?

3. Slurm Job Outcomes

- What are the frequencies and durations of Slurm jobs by outcome?
- Is job outcome affected by the total number of ended or successfully completed jobs?

4. Slurm Partition Usage Breakdown

5. Slurm Unresponsiveness Analysis

Frequencies of Job Outcomes: Number of Successful Job Completions per Hour is Higher Than Other Job Outcomes

Descriptive Statistics For Different Job Outcomes

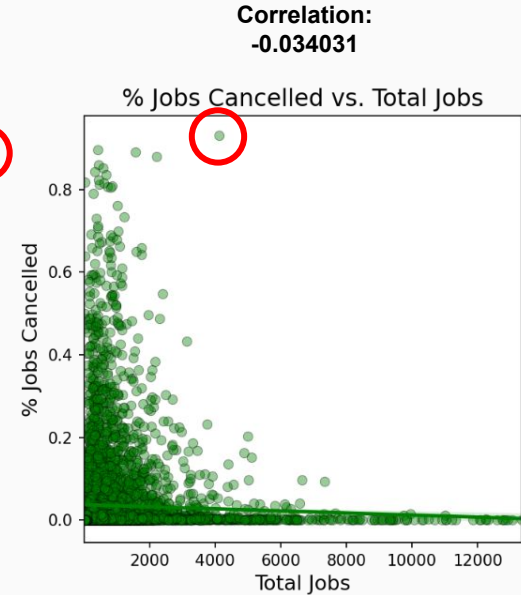
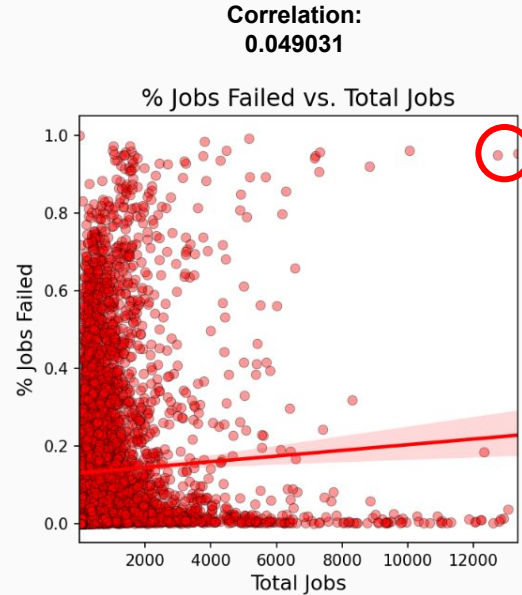
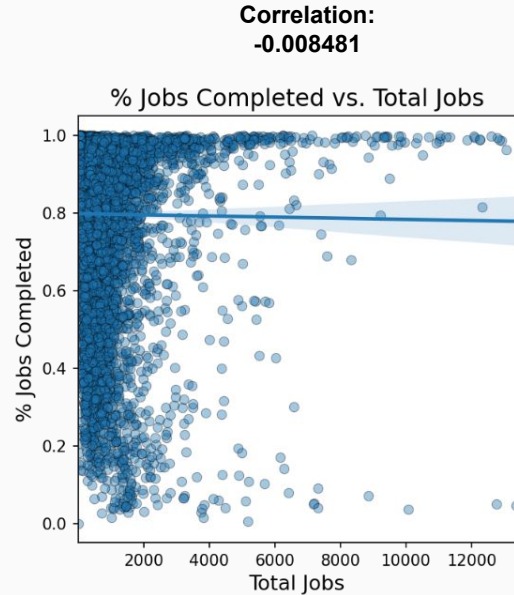
	Successfully Completed (Exit code 0:0 and completed state)	Failed (Various exit codes and failed state)	Cancelled (Various exit codes and cancelled state)
Mean per hour	2,213	1,150	215
Median per hour	1,062	382	102
Min	1	1	1
Max	13,784	15,523	3,870

Job Lengths by Outcome: Failed Jobs Generally Have Shorter Job Lengths

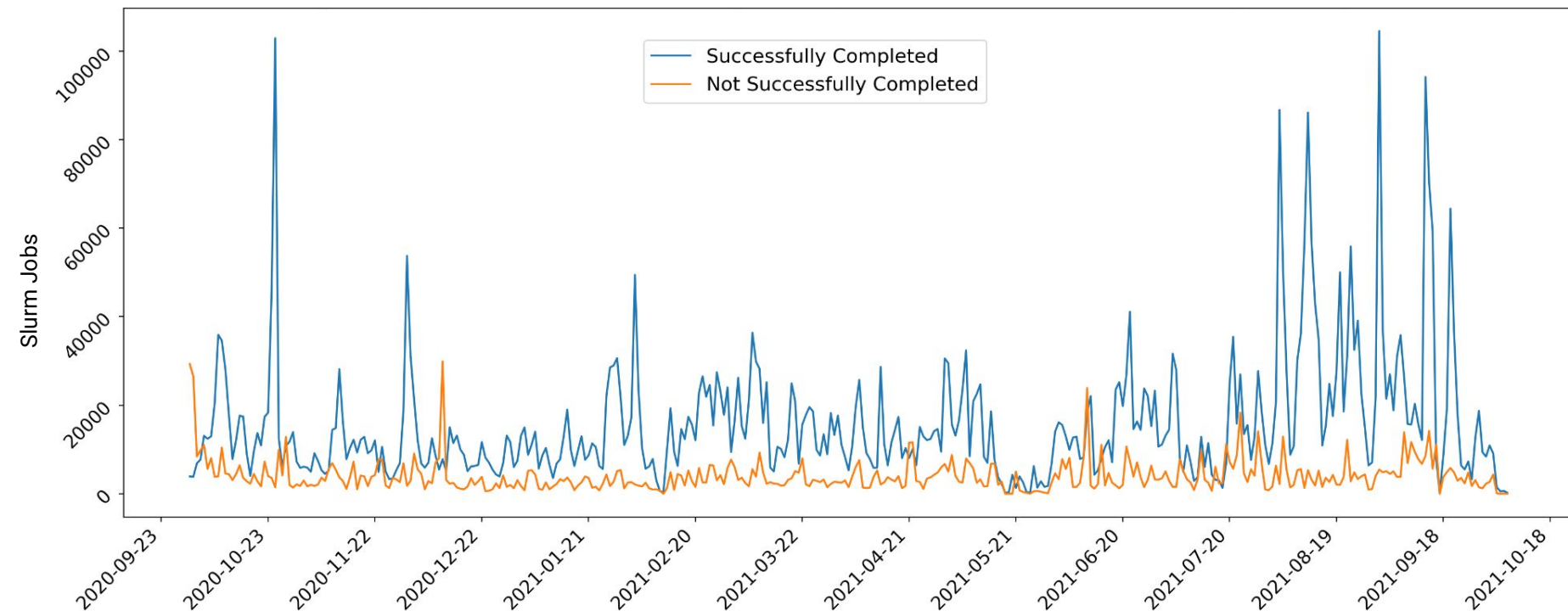
Descriptive Statistics For Job Length By Job Outcome

	Successfully Completed (Exit code 0:0 and completed state)	Failed (Various exit codes and failed state)	Cancelled (Various exit codes and cancelled state)
Mean job length	3 hours 31 min	1 hour 34 min	6 hours 15 min
Median job length	15min	1min	17min
Min	-1 days 23 hours 2 min	-1 days 23 hours 20 min	0 min
Max	22 days 23 hours 14 min	34 days 23 hours 15 min	24 days 22 hours 54 min

Job Outcomes vs Total Jobs: Total Number of Jobs Has No Effect on Job Outcomes



Successful vs Unsuccessful Jobs: Number of Successful Job Completions Do Not Impact Other Jobs



Agenda

1. Introduction

2. Slurm Usage Trends

- How many jobs does the Slurm scheduler process on a monthly/weekly/daily/hourly basis?

3. Slurm Job Outcomes

- What are the frequencies and durations of Slurm jobs by outcome?
- Is job outcome affected by the total number of ended or successfully completed jobs?

4. Slurm Partition Usage Breakdown

- What proportion of jobs are submitted to each Slurm partition?
- How do job outcomes compare by partition?

5. Slurm Unresponsiveness Analysis

Slurm Jobs by Partition: production Partition is the Most Used Partition

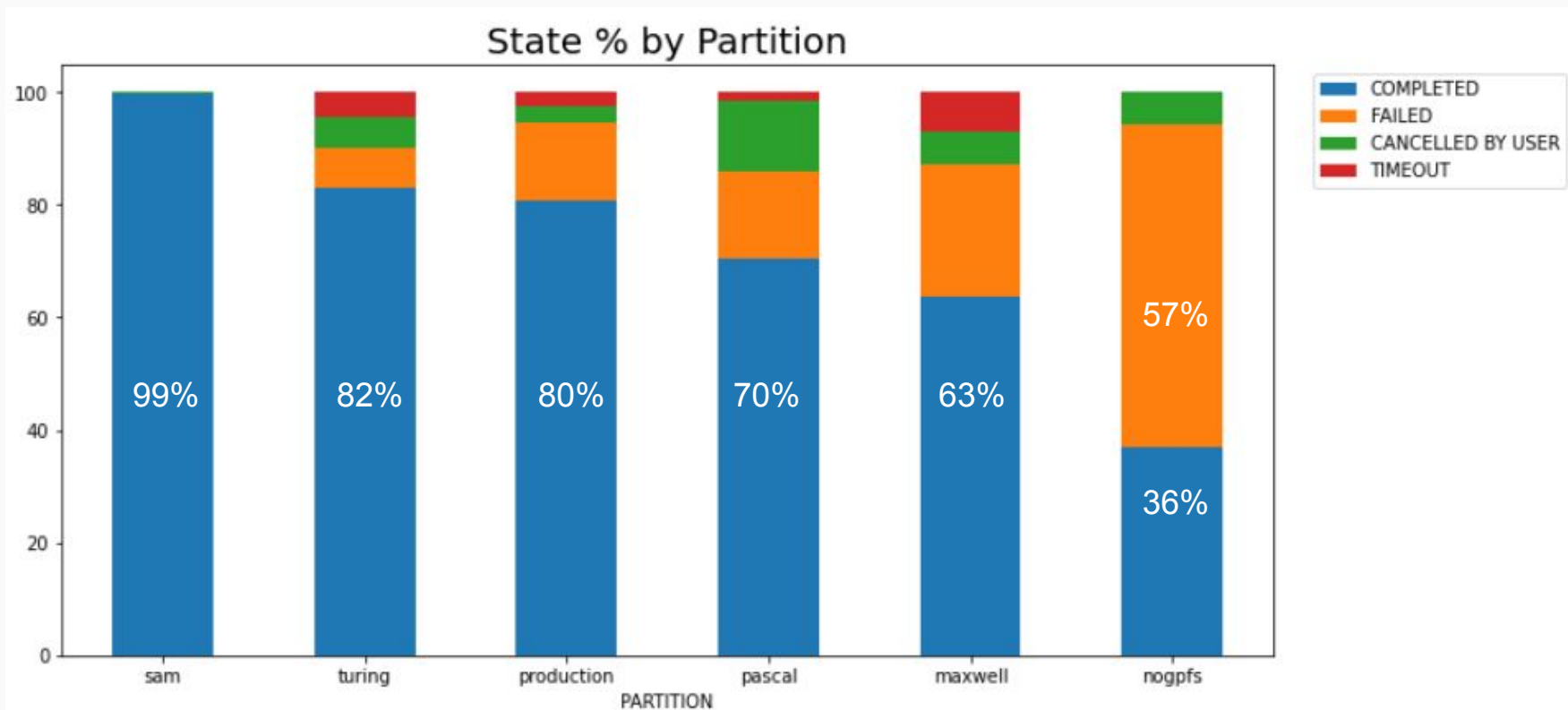
Proportion of Jobs by Partition

	PARTITION	COUNT	PROPORTION
1	production	7019136	95%
2	nogpfs	147044	2%
3	pascal	124451	2%
4	sam	64967	1%
5	turing	21412	0%
6	maxwell	11278	0%
7	cgw-maizie	4308	0%
8	debug	1616	0%
9	cgw-platypus	376	0%
10	cgw-dsi-gw	228	0%
11	cgw-capra1	153	0%
12	cgw-dougherty1	125	0%
13	cgw-horus	61	0%
14	cgw-cqs1	28	0%
15	cgw-hanuman	21	0%
16	cgw-sideshowbob	14	0%
17	cgw-vm-qa-flatearth1	9	0%
18	cgw-tbi01	7	0%

Job Outcomes by Partition: production Partition

	STATE	COUNT	PROPORTION
1	COMPLETED	241942.088660	78%
2	FAILED	49475.584610	16%
3	CANCELLED by USER	9217.337402	3%
4	TIMEOUT	8572.655686	3%
5	PREEMPTED	1121.457314	0%
6	OUT_OF_MEMORY	296.471394	0%
7	REQUEUED	84.258503	0%
8	NODE_FAIL	11.566929	0%
9	RUNNING	11.160000	0%
10	CANCELLED	1.666667	0%

Job Outcomes by Partition: nogpfs Partition has Highest Percentage of FAILED jobs



Agenda

1. Introduction

2. Slurm Usage Trends

- How many jobs does the Slurm scheduler process on a monthly/weekly/daily/hourly basis?

3. Slurm Job Outcomes

- What are the frequencies and durations of Slurm jobs by outcome?
- Is job outcome affected by the total number of ended or successfully completed jobs?

4. Slurm Partition Usage Breakdown

- What proportion of jobs are submitted to each Slurm partition?
- How do job outcomes compare by partition?

5. Slurm Unresponsiveness Analysis

- How often and for how long is the Slurm scheduler unresponsive?
- What is causing the Slurm scheduler to become unresponsive?

Identifying Slurm Scheduler Unresponsiveness Using CE5/6 Server Logs

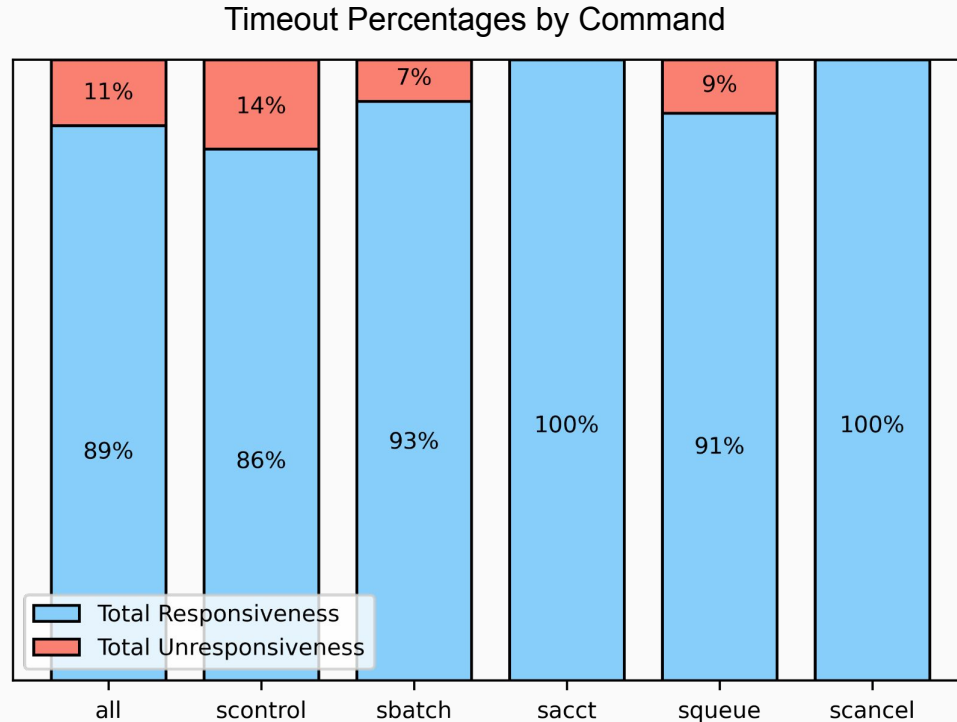
Servers CE5 and CE6 submit jobs to ACCRE's Slurm scheduler from the Open Science Grid.

Commands executed by CE5/6 provide insight into when the scheduler was responsive.

Example of Unresponsive Command (aka Timeout)

TIMESTAMP	2020-10-18 06:53:44
USER	9204
RETRY	0
RUNTIME	20.038464
RETURNCODE	1
COMMAND	sbatch

Timeout Percentages: Unresponsiveness Varies by Executed Command

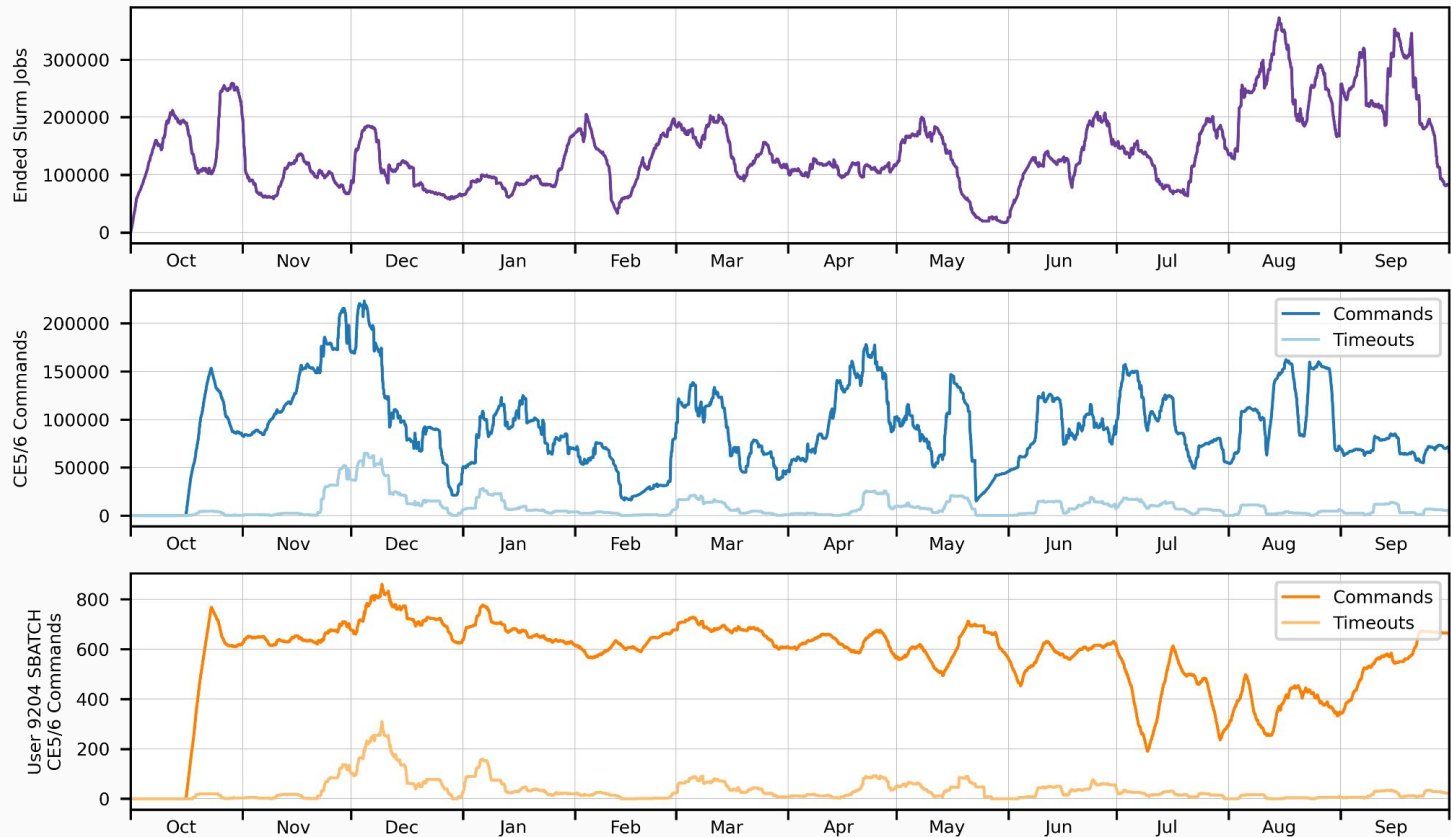


- SBATCH** Submit a batch script to Slurm
- SCONTROL** View or modify Slurm configuration and state
- SACCT** Displays accounting data for all jobs and job steps in the Slurm job accounting log or Slurm database
- SQUEUE** View information about jobs located in the Slurm scheduling queue
- SCANCEL** Used to signal jobs or job steps that are under the control of Slurm

Measuring Unresponsiveness: Timeout Statistics In Seconds

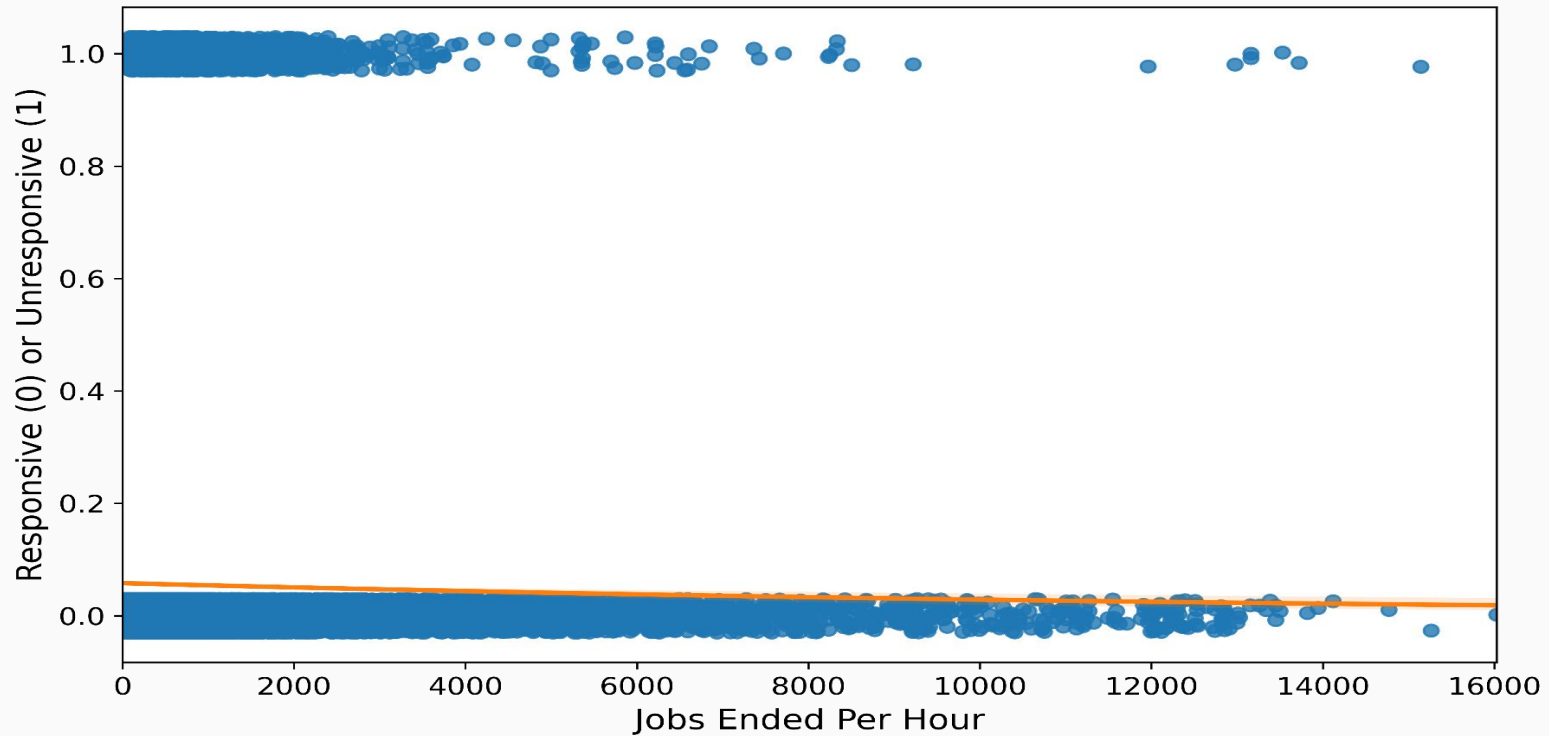
	<u>Each Timeout</u>	<u>Consecutive</u>	<u>Non-Consecutive</u>
Mean	21.64	31.98	21.24
Standard Deviation	5.53	96.76	4.93
Minimum	15.00	15.00	15.05
25th Percentile	20.05	20.04	20.05
Median	20.16	23.87	20.17
75th Percentile	20.42	40.03	20.53
Maximum	137.44	10991.21	137.44

Hypothesis: Lots of Jobs Ended in Short Time Period are Causing Unresponsiveness

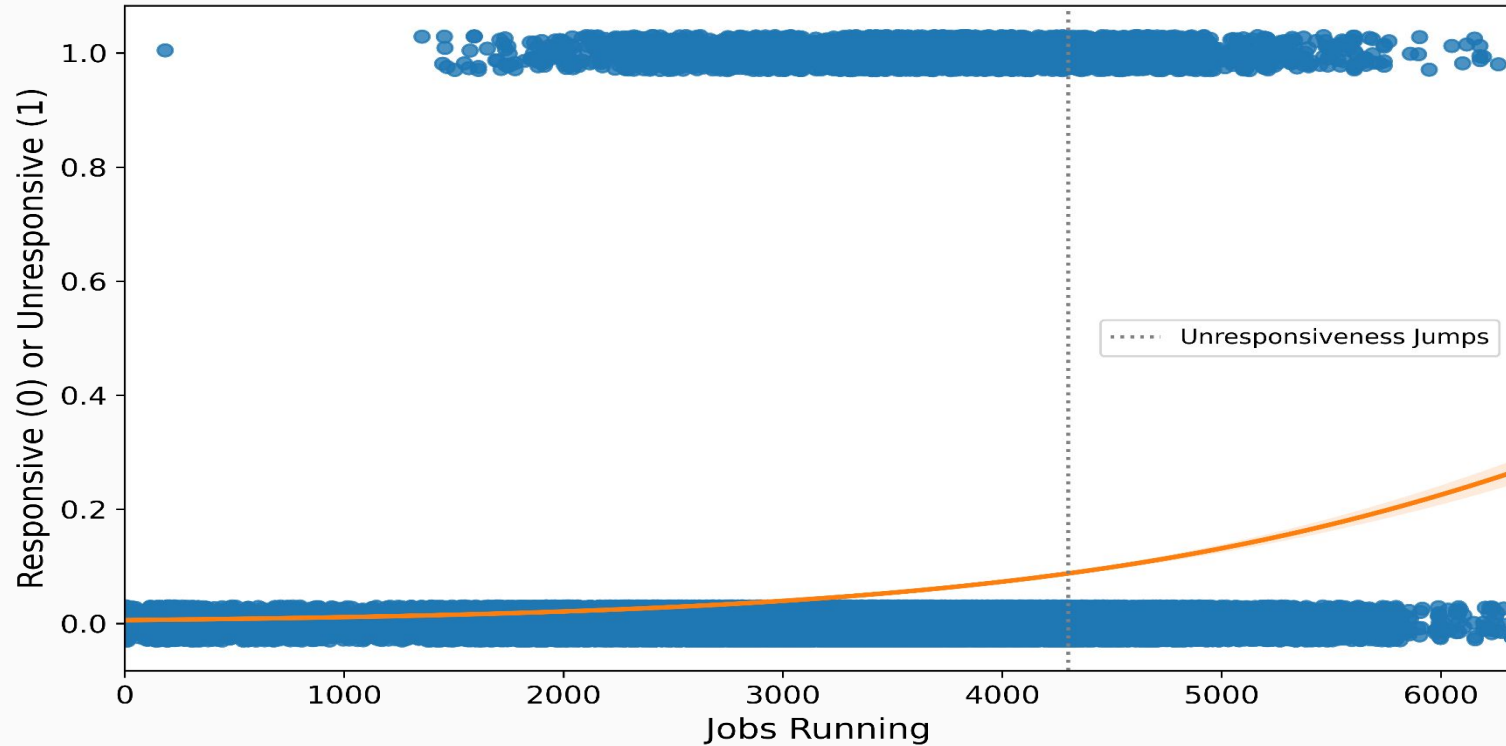


What is causing the Slurm scheduler to become unresponsive?

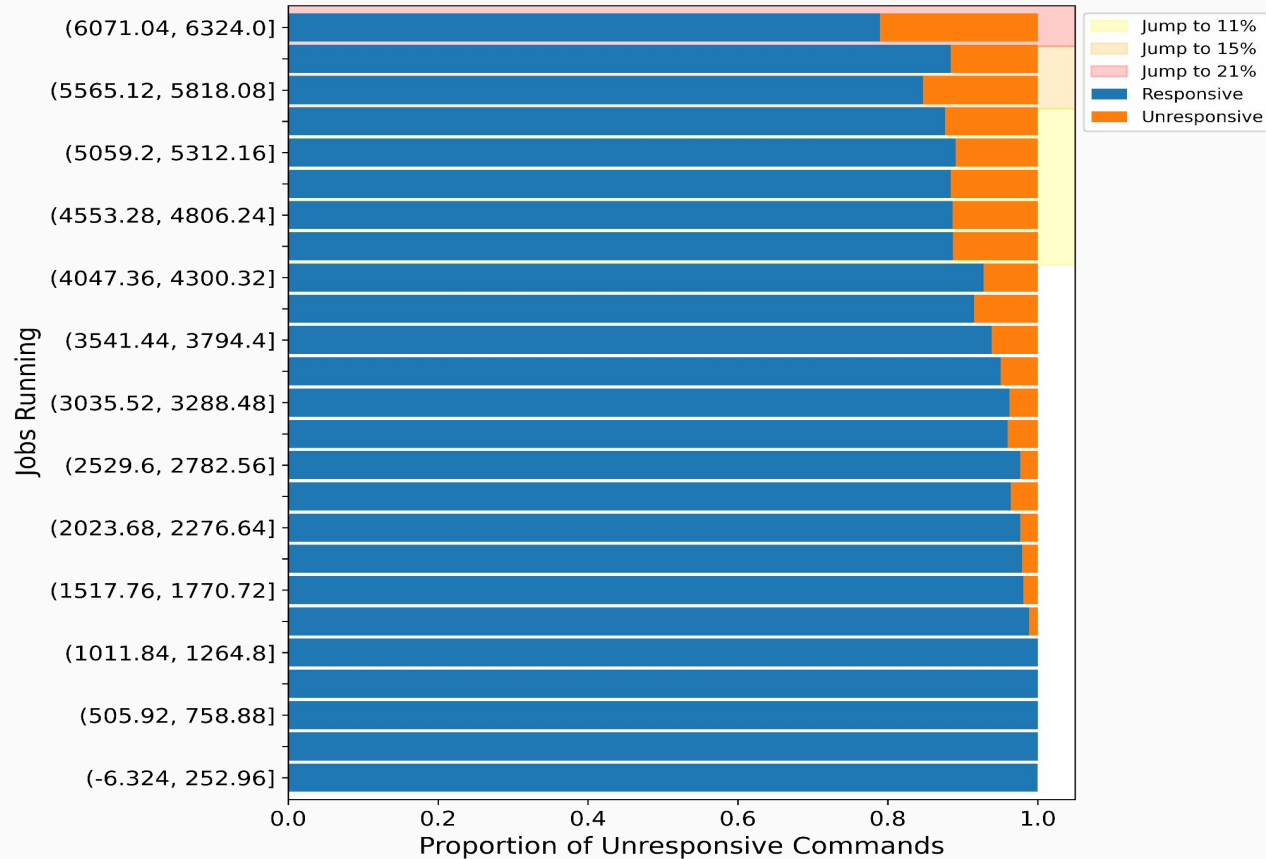
Unresponsiveness vs Jobs Ended: No Correlation or Negative Correlation



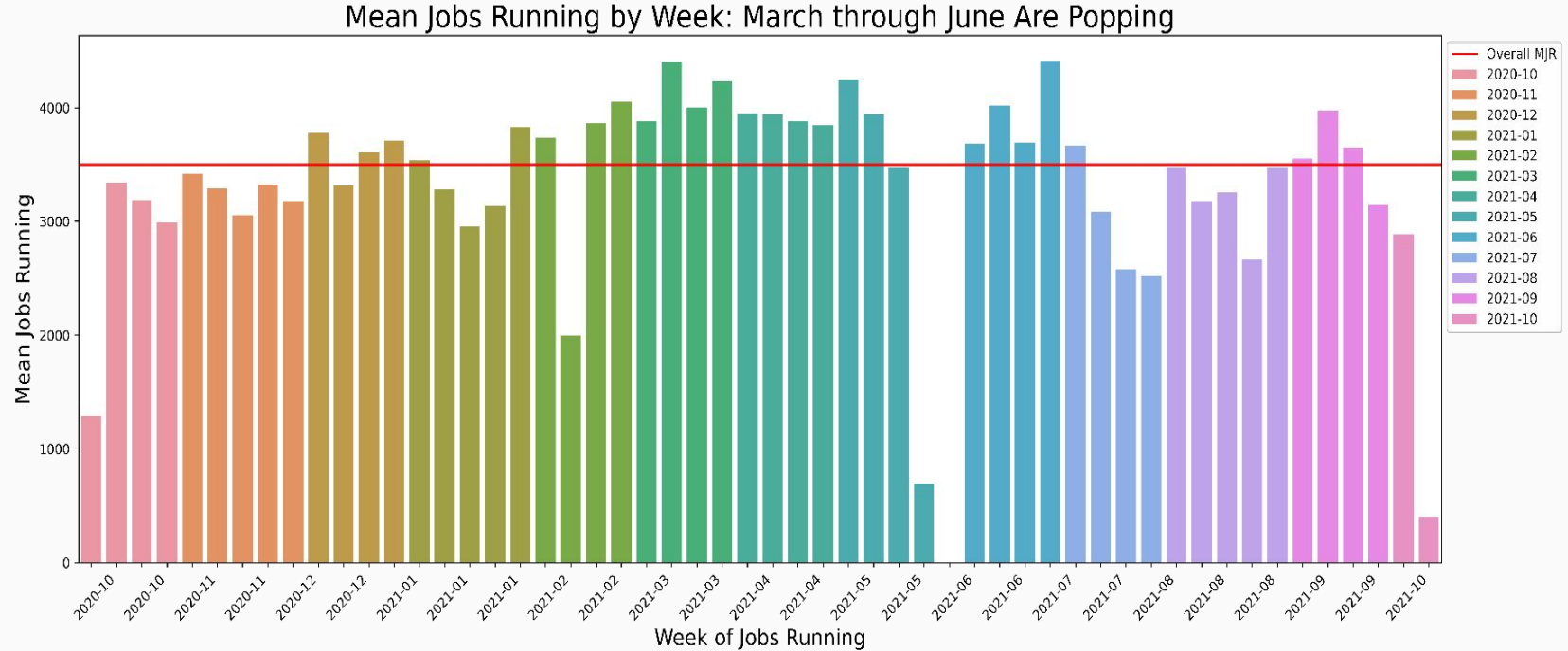
Unresponsiveness vs Jobs Running: Positive Correlation



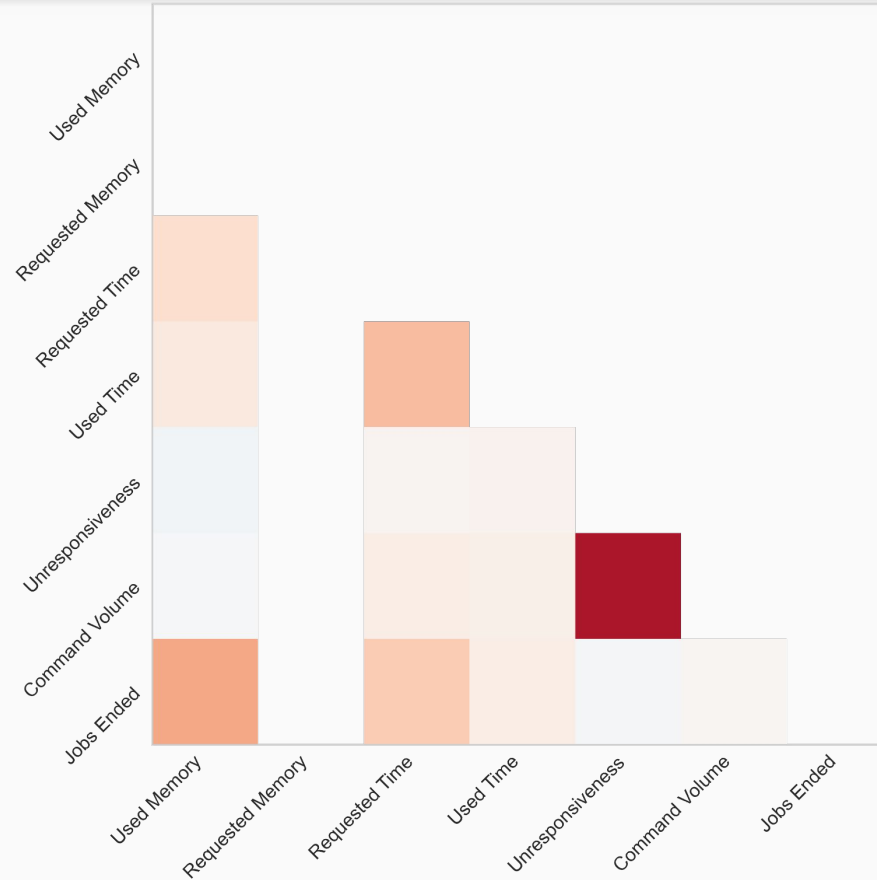
Unresponsiveness vs Jobs Running: Threshold Around 4,300 Jobs Running



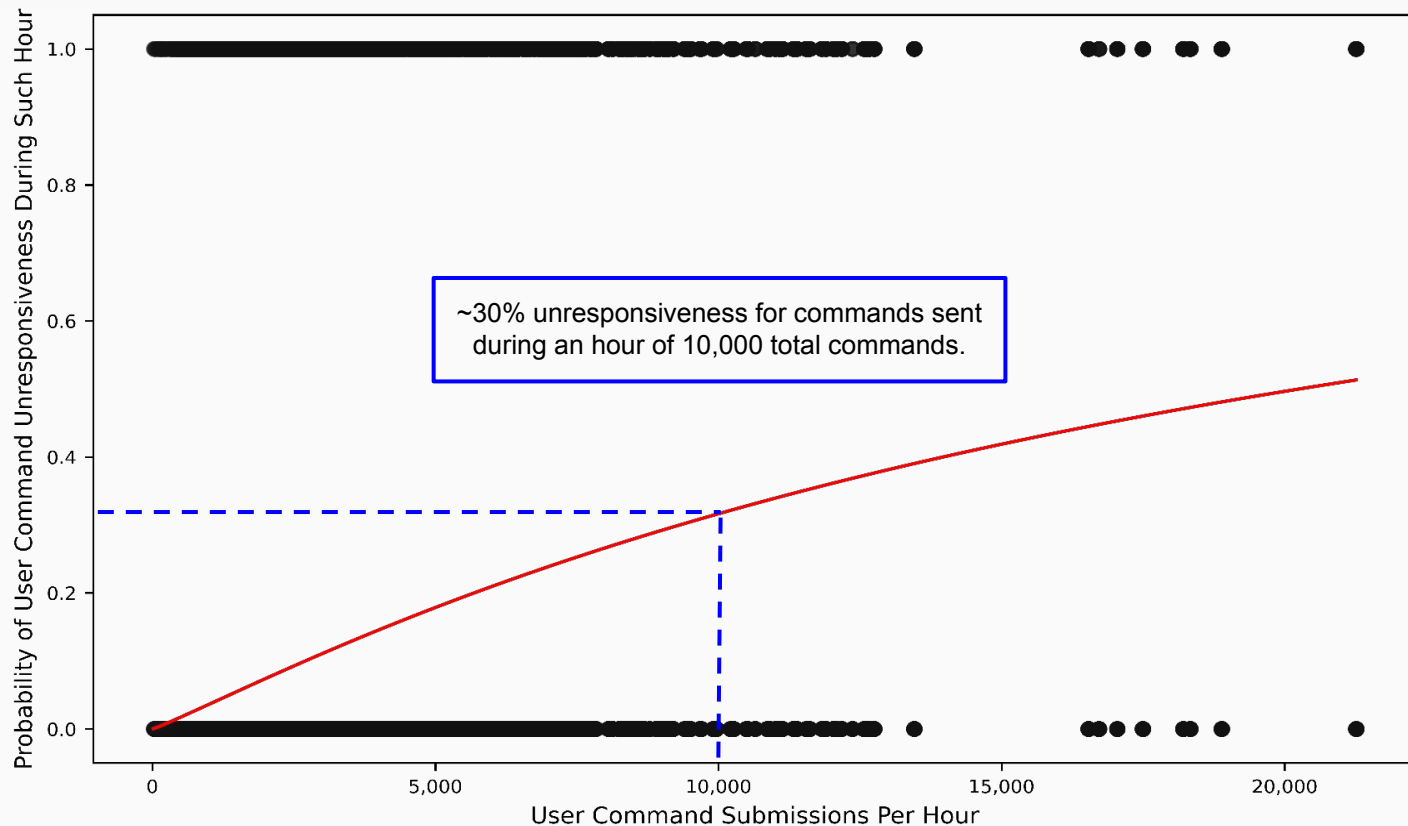
Takeaways About Increased Risk: Highest volume of jobs running looks seasonal



Alternative Hypothesis: Unresponsive Counts Caused by User Command Volume Based on Slurm Scheduler Data



Alternative Results: Increased Unresponsiveness During High Command Volume Periods in Slurm Scheduler Data



Closing Summary of Project Findings

- Most job completions occur in **August** and **September**, on **Tuesday**, during **late morning** to **evening**, and in the **production** partition. Job completions have increased over time.
- Most jobs complete **successfully**, use **only 25% of requested memory**, and last **only 2 hours**.
- Most **failed jobs** in the **maxwell** and **nogpfs** partitions. Most scheduler **unresponsiveness** with **scontrol**.
- The **number of jobs ended or completed** has a negative correlation with scheduler unresponsiveness.
- The **number of jobs concurrently running** has a positive correlation with scheduler unresponsiveness.
- The **number of user commands submitted on the scheduler** has a positive correlation with scheduler unresponsiveness.

Acknowledgements

Data Science 5 Cohort

Bryan Finlayson
James Gilbert
Christopher Harrelson
Ross Kimberlin
Eugenia Kolesova
Habeeb Kotun Jr.
Eli Lavender
Yvonne Martinez
Connor Merry
George Mills
Christopher Mulvey

Jacob Parks
Jessica Ra
Jake Randolph
Vahideh Rasekhi
Vamsi Reddy
Joshua Rio-Ross
Conrad Reihsmann
Neda Taherkhani
Rohit Venkat
Alex Zhang

Instructors

Michael Holloway
Mahesh Rao

Teaching Assistants

Veronica Ikeshoji-Orlati
Alvin Wendt

Community Partners

Eric Appelt

