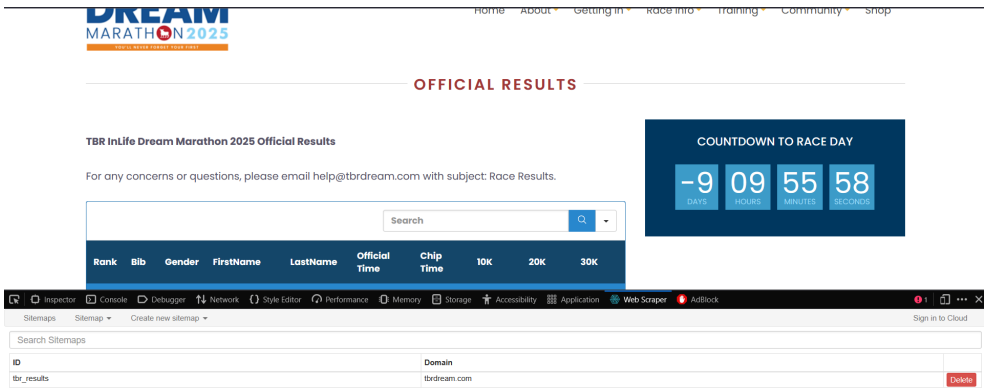
AWS Project: Visualizing Marathon Data Using AWS QuickSight

In this project, I used AWS QuickSight to analyze race results data from the recent Bull Runner Dream Marathon. AWS QuickSight is an AWS serverless Business Intelligence service that allows for quick and seamless Data Visualization.

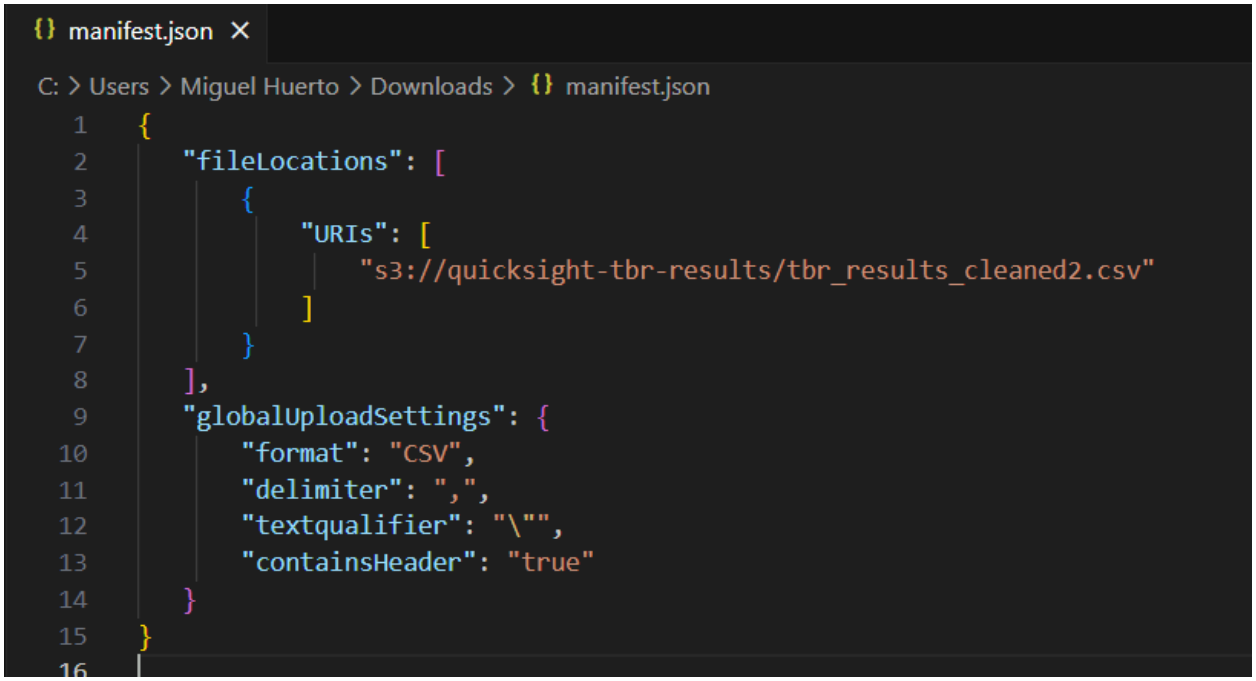
To gather the data, I downloaded WebScraper.io to scrape the data from the race results website. I created the “tbr_results” sitemap with two selectors, one for the main table, and one for the pagination.



Once all the setup was complete, I scraped the site automatically and gathered the results in a CSV file. (Names and Bib Numbers were hidden for privacy)

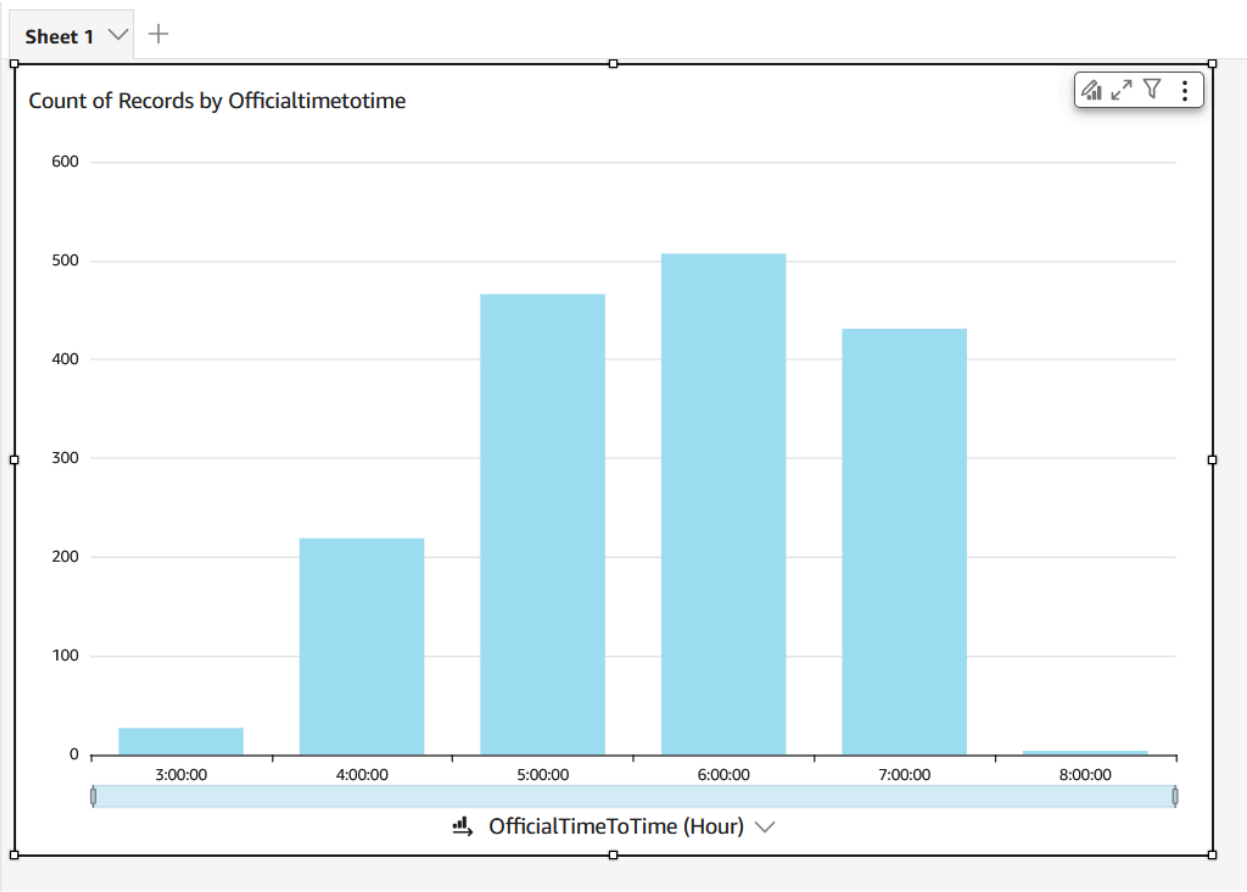
RunnerID	Rank	Bib	Gender	FirstName	LastName	OfficialTime	ChipTime	10K	20K	30K
			Male			7:49:31	7:48:23	1:45:07	1:50:08	2:01:54
			Male			3:20:12	3:20:12	0:46:16	0:45:35	0:49:58
			Male			3:31:46	3:27:43	0:55:33	0:51:18	0:50:36
			Male			3:33:59	3:33:55	0:49:25	0:50:23	0:54:32
			Male			3:44:59	3:44:55	0:55:25	0:53:16	0:56:46
			Male			3:46:26	3:46:24	0:55:21	0:54:32	0:56:23
			Male			3:47:01	3:46:40	0:55:05	0:53:53	0:55:19
			Male			3:48:59	3:46:38	0:57:45	0:54:37	0:55:56
			Female			3:50:51	3:50:07	0:51:00	0:52:00	0:57:10
			Male			3:51:24	3:51:16	0:51:32	0:54:37	1:00:17
			Male			3:52:12	3:51:02	0:55:02	0:54:31	0:58:37
			Female			3:52:30	3:52:13	0:55:34	0:57:59	0:58:20
			Male			3:52:38	3:51:48	0:51:24	0:54:43	0:59:58
			Male			3:52:47	3:52:15	0:57:47	0:56:24	0:57:35
			Male			3:52:59	3:52:52	0:55:26	0:53:14	0:51:52
			Female			3:53:10	3:53:09	0:55:33	0:58:12	0:58:42
			Male			3:54:26	3:54:11	0:56:22	0:54:10	1:01:22
			Male			3:55:14	3:54:46	0:54:07	0:55:29	1:01:12
			Male			3:55:56	3:55:54	0:55:56	0:58:10	0:59:42
			Male			3:55:56	3:55:16	0:55:12	0:57:56	1:00:52

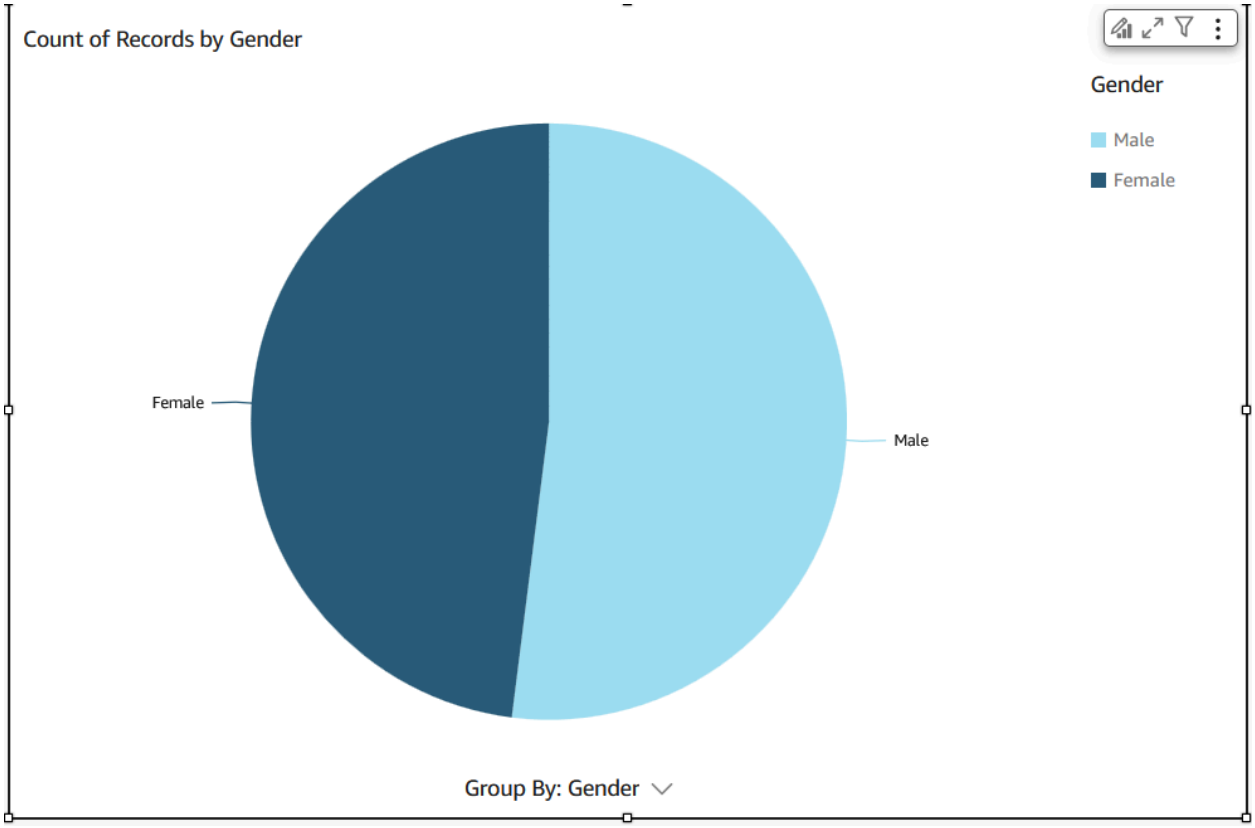
Next, the manifest.JSON file was created to indicate to quicksight where the data is stored alongside directives on how to read the CSV file dataset.

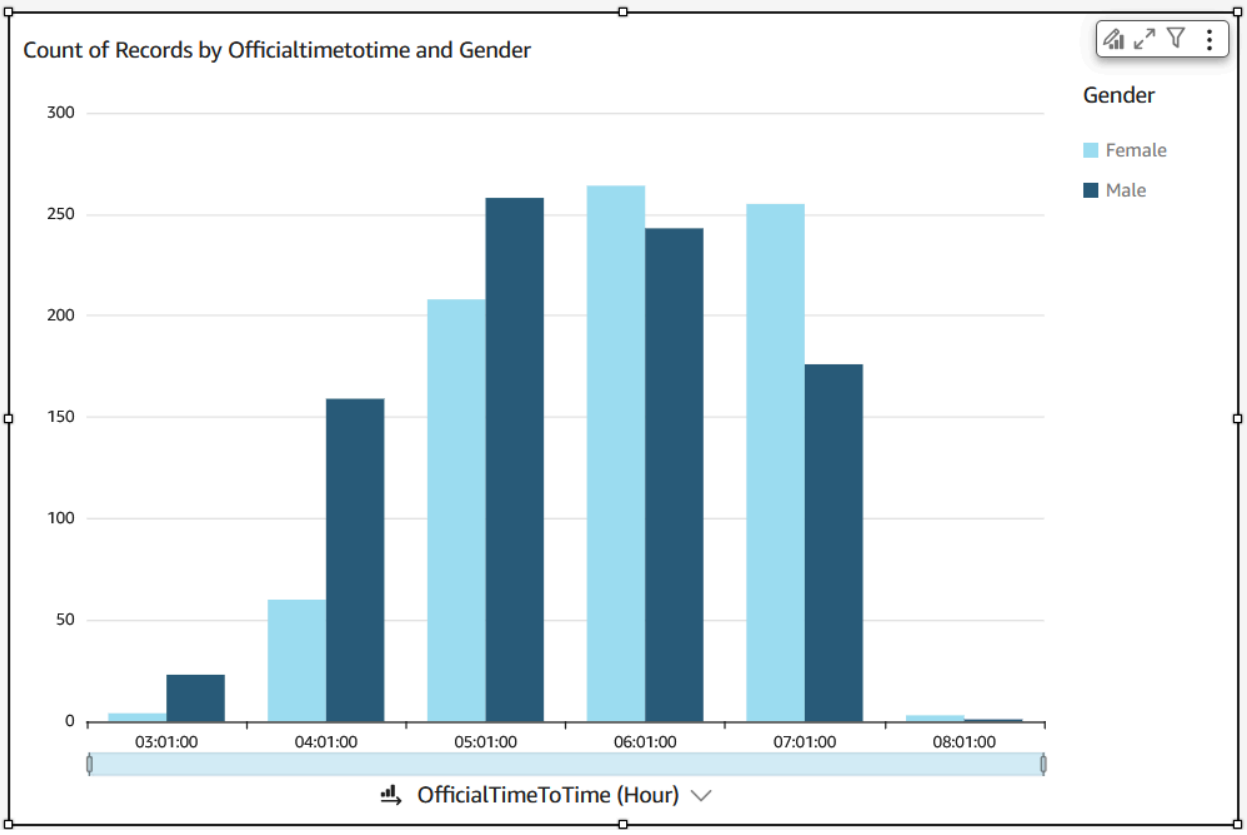


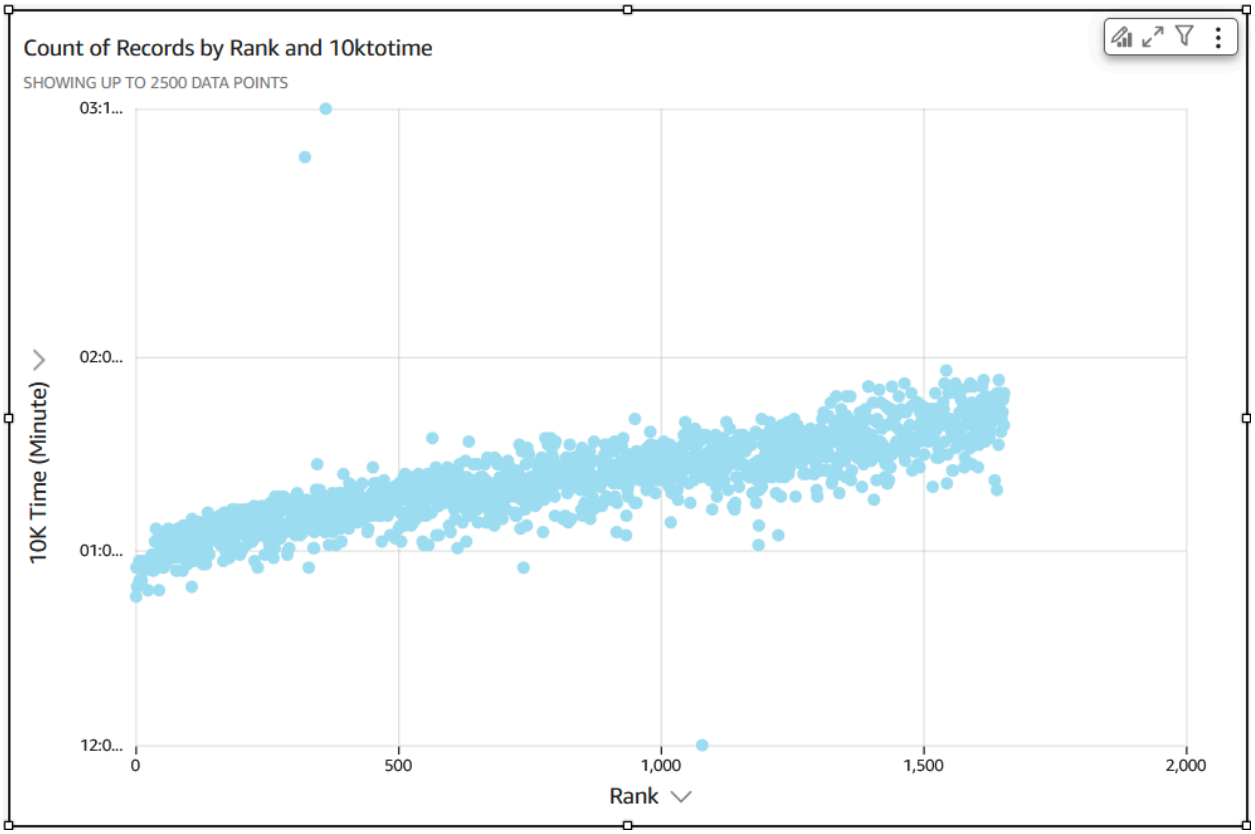
I created an S3 bucket to store both the dataset and the manifest.JSON file. This is the bucket where QuickSight will be accessing the data. On AWS QuickSight, once access is granted to the bucket, visualizations can now be made. Since QuickSight interprets the times as String, I had to create a Calculated Field to convert string into DateTime before using them for visualizations.

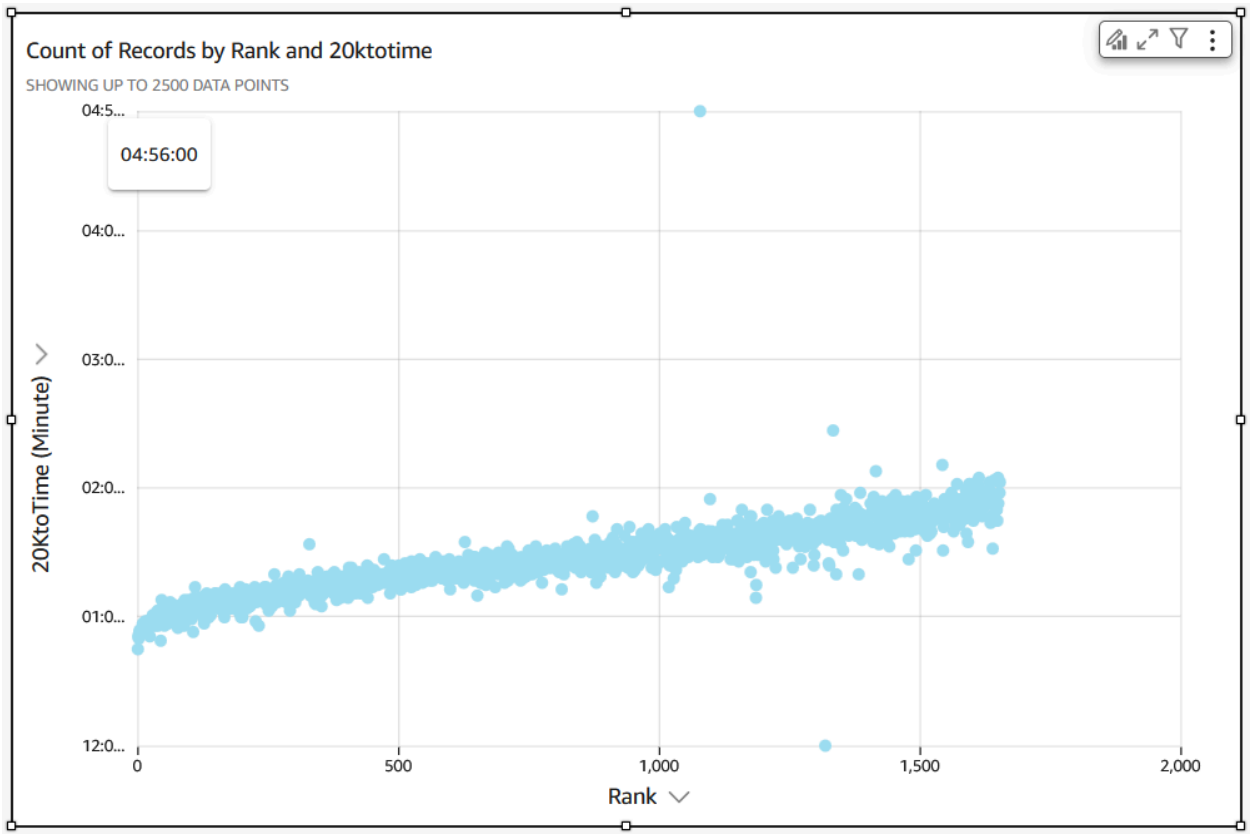
Graph 1: Count of Records by Official Time

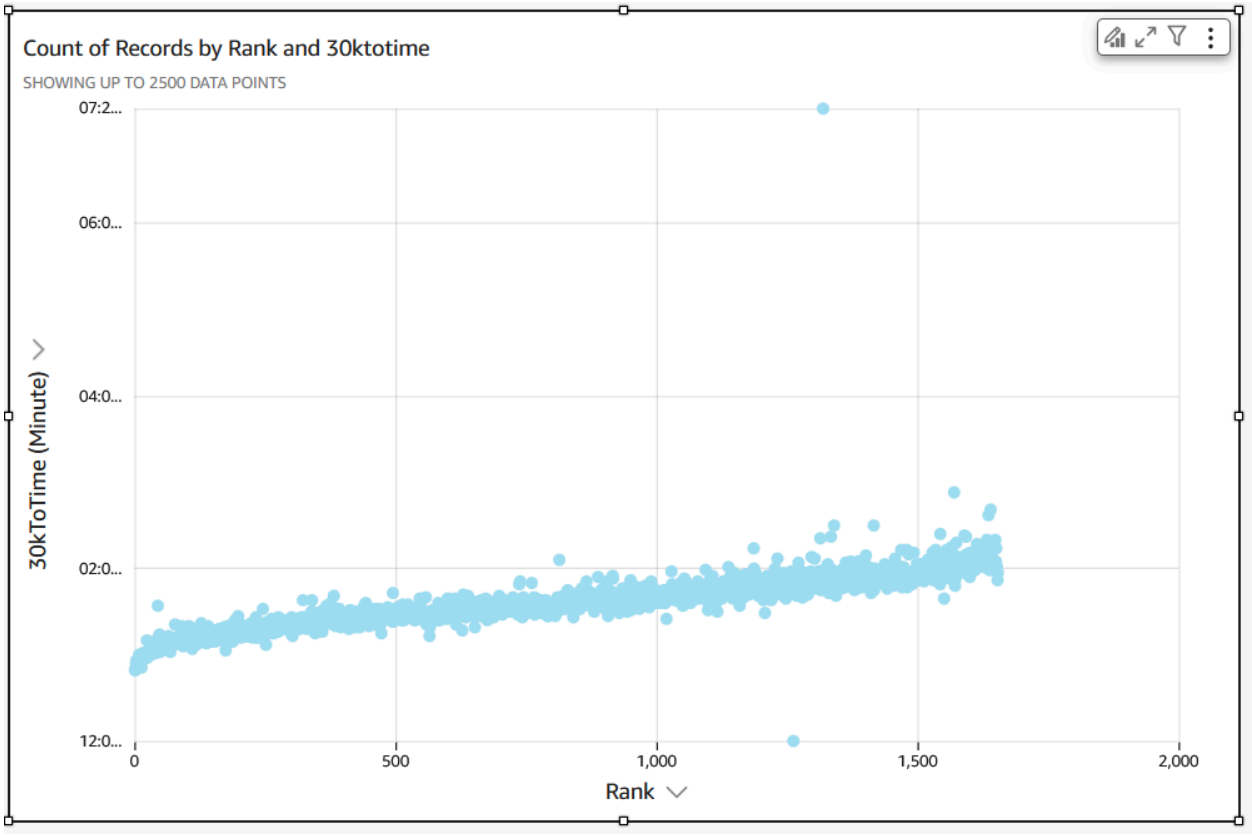



Count of Records by Gender
Gender
Male
Female
Female
Male
Group By: Gender


Count of Records by Officialtimetotime and Gender
Gender
Female
Male
300
250
200
150
100
50
0
03:01:00
04:01:00
05:01:00
06:01:00
07:01:00
08:01:00
OfficialTimeToTime (Hour)


Count of Records by Rank and 10ktotime
SHOWING UP TO 2500 DATA POINTS
03:1...
02:0...
01:0...
12:0...
10K Time (Minute)
0
500
1,000
1,500
2,000
Rank


Count of Records by Rank and 20ktotime
SHOWING UP TO 2500 DATA POINTS
04:5...
04:56:00
04:0...
03:0...
02:0...
01:0...
12:0...
20KtoTime (Minute)
0
500
1,000
1,500
2,000
Rank


Count of Records by Rank and 30ktotime
SHOWING UP TO 2500 DATA POINTS
07:2...
06:0...
04:0...
02:0...
12:0...
30kToTime (Minute)
0
500
1,000
1,500
2,000
Rank