



Air Quality Visuals With MongoDB

MongoDB tool exploration

Kevin Liao
Big Data Tools & Techniques
2024



Goals

- Explore Visuals with MongoDB
- Get more comfortable with cloud based utilities instead of self hosted tools
 - Technology constantly changes, should learn on IaaS/SaaS more
- Not spend money



Steps

1. Obtain Data:
<https://www.kaggle.com/datasets/threnjen/40-years-of-air-quality-index-from-the-epa-daily/data>
2. Cut/Clean Data
 - a. Too large for Free Atlas (reduce data)
 - b. `awk -F ',' '($2 >= "1980-01-01" && $2 <= "2011-12-31") {next} {print} '`
`aqi_daily_1980_to_2021.csv > aqi_daily_2011_to_2021.csv`
3. Register a MongoAtlas account
 - a. <https://cloud.mongodb.com>
4. Download MongoDB Compass (Client Interface)
 - a. <https://www.mongodb.com/try/download/compass>
5. Connect Compass with Atlas
 - a. Configure Compass to connect to your MongoDB Atlas cluster using your cluster credentials.
6. Familiarize with MongoDB Atlas Charts



Data

Content

This set includes daily reports of air quality index from various US Metro areas, as well as geographic data for the collection locations. This set has over 10 million rows.

Want a smaller set? I also have ANNUAL metrics available [here](#). The yearly dataset is ~34k rows.

Acknowledgements

Data sourced directly from the [Environmental Protection Agency](#)

Banner: Photo by [Maxim Tolchinskiy](#) on [Unsplash](#)

```
_id: ObjectId('662d5f6d15f4125655a7e794')
State Name : "Alabama"
Date : 2021-01-01T00:00:00.000+00:00
AQI : 30
Category : "Good"
Defining Parameter : "Ozone"
Latitude : 34.289001
Longitude : -85.970065
County Name : "DeKalb"
```

Cut Data

- Limited to 500 Mb
- Had to cut daily data from 10 million records from 1980-2021 to 1.37 million records 2017-2021
- `awk -F ',' ' ($2 >= "1980-01-01" && $2 = "2016-12-31") {next} {print} ' $input_file > $output_file`

```
cheerio@cheerio-precision:~/Rowan_Local/BigData/BigDataFP/data/air_quality/air_quality (main)$ cat aqi_daily_2017_to_2021.csv | wc -l
1379828
```

```
cheerio@cheerio-precision:~/Rowan_Local/BigData/BigDataFP/data/air_quality/air_quality (main)$ cat aqi_daily_1980_to_2021.csv | wc -l
10158528
```

```
cheerio@cheerio-precision:~/Rowan_Local/BigData/BigDataFP/data/air_quality/air_quality (main)$ ll
total 926556
```

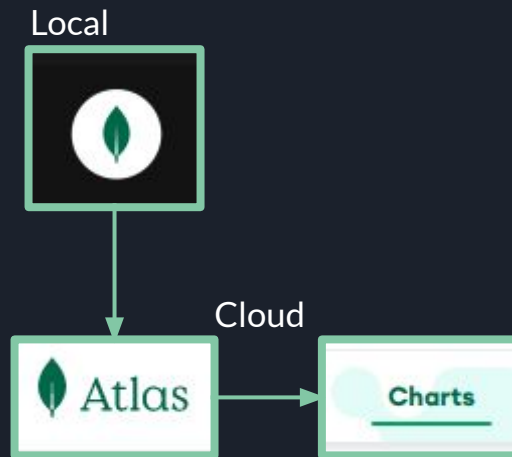
```
drwxrwxr-x 2 cheerio cheerio      4096 Apr 27 16:40 ./
drwxrwxr-x 3 cheerio cheerio      4096 Apr 27 15:45 ../
-rw-rw-r-- 1 cheerio cheerio 669029200 Aug 28  2021 aqi_daily_1980_to_2021.csv
-rw-rw-r-- 1 cheerio cheerio 190672791 Apr 27 15:40 aqi_daily_2011_to_2021.csv
-rw-rw-r-- 1 cheerio cheerio  89074250 Apr 27 16:14 aqi_daily_2017_to_2021.csv
```

```
)$ ls
air_quality.zip      aqi_daily_1980_to_2021.csv
air_quality.zip.orig aqi_daily_2017_to_2021.csv
cheerio@cheerio-precision:~/Rowan_Local/BigData/BigDataFP/data/air_quality
```

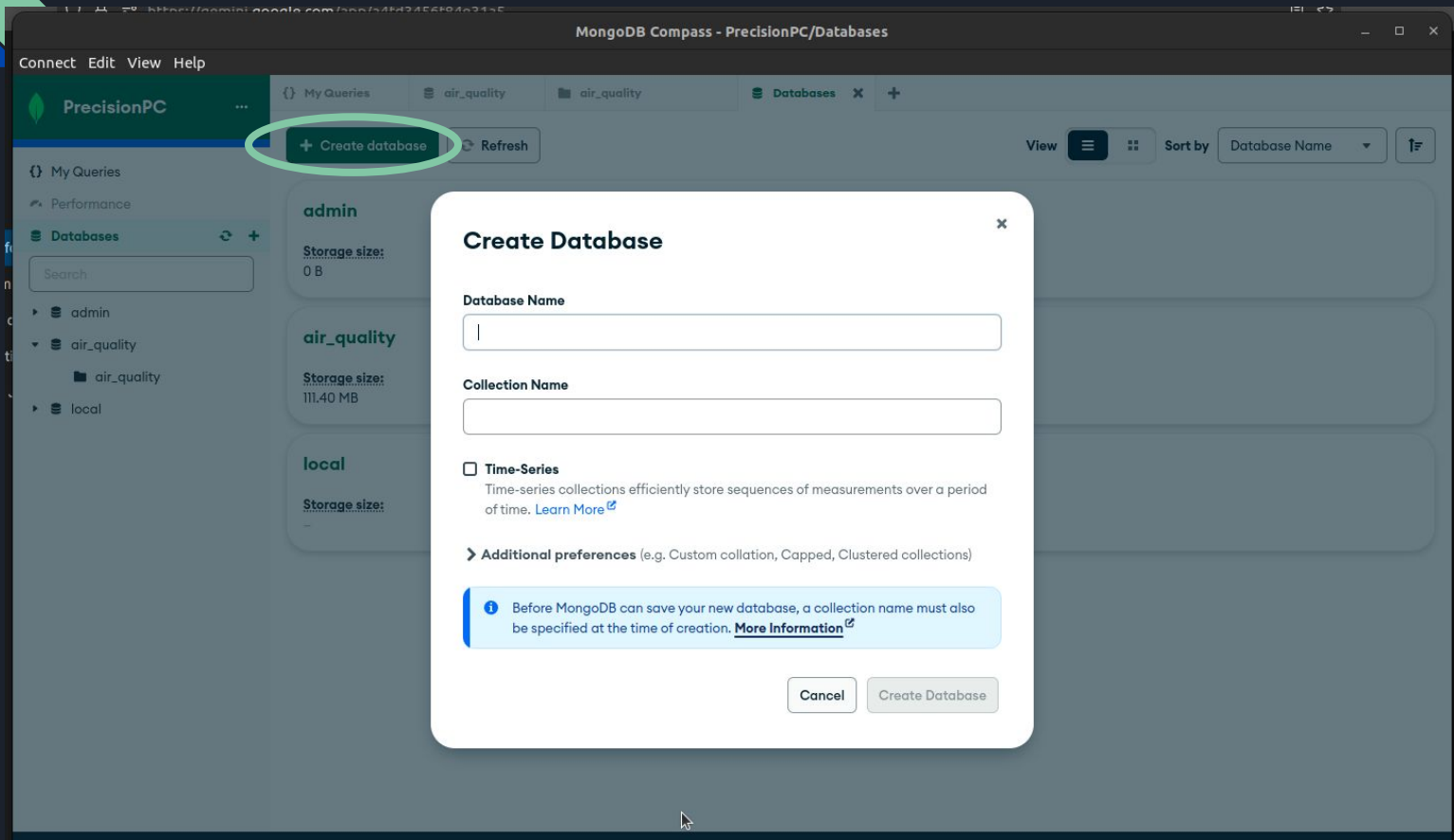
```
)$ cat aqi_daily_2017_to_2021.csv | egrep -v "2021" | wc -l
1339545
```

MongoDB

- MongoDB Compass
 - Mongo's GUI Desktop Utility
- MongoDB Atlas
 - Mongo's online SaaS
 - Billing structure like Azure
- MongoDB Charts
 - Online SaaS visualization tools like Tableau
 - Free tier limits



Import Data w/ MongoDB Compass



Import Data w/ MongoDB Compass

The screenshot displays the MongoDB Compass interface for the 'air_quality' collection. The left sidebar shows the database structure with 'air_quality' selected. The main panel shows the 'Documents' tab with a list of documents. A green arrow points from the 'ADD DATA' button to a summary overlay on the right. The overlay displays the following information:

- Documents: 1.4m
- Storage Size: 49.3MB
- Avg. Size: 182B


The document viewer shows the following JSON document:


```
{
  "_id": "ObjectId('662d5f6d15f4125655a7e796')",
  "State Name": "Alabama",
  "Date": "2021-01-03T00:00:00.000+00:00",
  "AQI": 34,
  "Category": "Good",
  "Defining Parameter": "Ozone",
  "Latitude": 34.289001,
  "Longitude": -85.970065,
  "County Name": "DeKalb"
}
```



MongoDB Atlas Free Tier

Shared Cluster

Pay-as-you-go! Clusters are billed hourly with monthly invoices.

 Amazon Web Services

 Microsoft Azure

 Google Cloud

Cluster Tier	Storage	RAM	vCPUs	Base Price
M0	512 MB	Shared	Shared	Free forever
M2	2 GB	Shared	Shared	\$9/mo
M5	5 GB	Shared	Shared	\$25/mo

options available



Shared

from \$0/month

Try for Free

(i) Free forever for free clusters

For learning and exploring MongoDB in a cloud environment. Basic configuration options.

- ✓ 512MB to 5GB of storage
- ✓ Shared RAM
- ✓ Upgrade to dedicated clusters for full functionality
- ✓ No credit card required to start

[View pricing](#)

<https://www.mongodb.com/pricing>

MongoDB Atlas Free Tier

MongoDB Version and Storage Engine	Atlas uses MongoDB 6.0 for M0/M2/M5 clusters.
MongoDB Version Upgrade	You can't upgrade the MongoDB version that M0/M2/M5 clusters run. Atlas upgrades M0 free clusters or M2/M5 shared clusters to the newest MongoDB version after several patch versions become available for that version. To learn more, see MongoDB Versioning .
Cluster Tier	<p>You must select the M0 cluster tier to deploy a free cluster.</p> <div><p>NOTE</p><p>You can deploy only one M0 cluster per project.</p></div>
Cluster Memory	You can't configure memory for M0 free clusters or M2/M5 shared clusters.
Cluster Storage	You can't configure storage size for M0 free clusters or M2/M5 shared clusters.
Replication Factor	Replication Factor is set to 3 Nodes and you can't modify it for M0 free clusters or M2/M5 shared clusters.
Replica Set Tags	M0 free clusters and M2/M5 shared clusters don't have pre-defined replica set tags .
Do You Want A Sharded Cluster	You can't deploy a M0 free cluster or M2/M5 shared cluster as a Sharded Cluster.
Do You Want To Enable Backup	You can't enable backups on M0 free clusters.

<https://www.mongodb.com/docs/atlas/reference/free-shared-limitations/>

MongoDB Atlas

The screenshot displays the MongoDB Atlas web interface. The browser address bar shows the URL: `https://cloud.mongodb.com/v2/6534227fee1db33d6c71c0d6#/overview`. The top navigation bar includes the Atlas logo, the organization name "Peyton's Org...", and links for "Access Manager" and "Billing". On the right, there are links for "All Clusters", "Get Help", and the user profile "Peyton".

The main navigation sidebar on the left lists various sections: Overview (selected), DEPLOYMENT, Database, Data Lake, SERVICES, Device Sync, Triggers, Data API, Data Federation, Atlas Search, Stream Processing, Migration, SECURITY, Backup, Database Access, Network Access, Advanced, New On Atlas (with a badge of 2), and Goto.

The "Overview" section for "PEYTON'S ORG - 2023-10-21 > PROJECT 0" features a "Database Deployments" card. This card includes a "Create deployment" button and a "Cluster0" section. Within this section, the "Connect" button is highlighted with a green circle, alongside "Upgrade" and "Edit" buttons. The "Data Size" is listed as "307.81 MB". Below these buttons are two action cards: "Browse collections" and "View monitoring", each with a right-pointing arrow. A "+ Add Tag" button is also present.

Below the Database Deployments card is the "Application Development" section, which includes a "Connect new" button. It contains a card titled "OPTIMIZE YOUR CONNECTION POOL" with the instruction "Use one MongoClient instance per application." Below this, there are tabs for "Drivers" and "AI Frameworks" (which is currently selected). A note states: "Enhance your Python application with popular Python AI Tech Stacks."

At the bottom, there is a section titled "DEVELOP INTELLIGENT SEARCH & CHAT WITH VECTOR SEARCH" with the instruction "Leverage Vector Search and the popular LangChain+HuggingFace stack. [Install PyMongo first.](#)"

On the right side of the interface, there are two panels. The "Toolbar" panel contains a "Featured Resources" section with links to PYMONGO resources: "Aggregations in Python", "Data Encryption Quickstart", "PyMongoArrow: Data Analysis", "Beanie ODM Quickstart", and "More Python Content". Below this is a "Sample Apps" section with links to "Flask Stack", "FAST API", and "FARM Stack". The "New On Atlas" panel shows a "2 NEW" badge and a link to "Learn about the latest feature enhancements on Atlas."

Connecting Compass to Atlas

The screenshot displays the MongoDB Atlas interface with a modal dialog titled "Connect to Cluster0". The dialog is divided into three steps: 1. Set up connection security (marked with a green check), 2. Choose a connection method (the current step), and 3. Connect. Under the heading "Connect to your application", several options are listed, each with an icon and a description:

- Drivers**: Access your Atlas data using MongoDB's native drivers (e.g. Node.js, Go, etc.)
- Data Explorer**: Browse your Atlas collections without leaving the UI
- Compass**: Explore, modify, and visualize your data with MongoDB's GUI (this option is highlighted with a green circle)
- Shell**: Quickly add & update data using MongoDB's Javascript command-line interface
- MongoDB for VS Code**: Work with your data in MongoDB directly from your VS Code environment
- Atlas SQL**: Easily connect SQL tools to Atlas for data analysis and visualization

At the bottom of the dialog are "Go Back" and "Close" buttons. The background shows the Atlas "Overview" page for "Project 0", with sections for "Database Deployments", "Browse collections", and "Application Development".

Connecting Compass to Atlas

Atlas Peyton's Org... Access Manager Billing

Project 0 Data Services App Services Charts

Overview PEYTON'S ORG - 2023-10-21 > PROJECT 0

Database Deployments Cluster0

Connect Upgrade Edit

Browse collections + Add Tag

Application Development OPTIMIZE YOUR CONNECTION POOL Use one MongoClient instance per application.

Drivers AI Frameworks

Enhance your Python application with popular Python AI Tech Stacks.

DEVELOP INTELLIGENT SEARCH & CHAT WITH VECTOR SEARCH Leverage Vector Search and the popular LangChain+HuggingFace stack. Install PyMongo first.

Connect to Cluster0

Set up connection security Choose a connection method Connect

Connecting with MongoDB Compass

I don't have MongoDB Compass installed I have MongoDB Compass installed

1. Choose your version of Compass

1.12 or later

See your Compass version in "About Compass"

2. Copy the connection string, then open MongoDB Compass

`mongodb+srv://cheerio:<password>@cluster0.so7kree.mongodb.net/`

Replace `<password>` with the password for the `cheerio` user. Ensure any options are URL encoded.

RESOURCES

Connect with Compass Import and Export Data

Access your Database Users Troubleshoot Connections

Go Back Close Review setup steps

Connecting Compass to Atlas

The screenshot shows the MongoDB Atlas console interface. A modal dialog titled "Connect to Cluster0" is open, guiding the user through the connection process. The dialog has three steps: "Set up connection security" (completed), "Choose a connection method" (current step), and "Connect" (pending). In the "Choose a connection method" step, the user has selected "I have MongoDB Compass installed". The dialog then prompts the user to "Choose your version of Compass", with a dropdown menu set to "1.12 or later". Below this, it instructs the user to "Copy the connection string, then open MongoDB Compass". The connection string, `mongodb+srv://cheerio:<password>@Cluster0.so7kree.mongodb.net/`, is displayed and highlighted with a green oval. A note below the string says: "Replace <password> with the password for the **cheerio** user. Ensure any options are URL encoded." At the bottom of the dialog, there are links for "RESOURCES" including "Connect with Compass", "Access your Database Users", "Import and Export Data", and "Troubleshoot Connections". The dialog also includes "Go Back", "Close", and "Review setup steps" buttons.

Atlas Peyton's Org... Access Manager Billing

Project 0 Data Services App Services Charts

Overview PEYTON'S ORG - 2023-10-21 > PROJECT 0

Database Deployments

Cluster0

Connect Upgrade Edit

Browse collections

+ Add Tag

Application Development

OPTIMIZE YOUR CONNECTION POOL

Use one MongoClient instance per application.

Drivers AI Frameworks

Enhance your Python application with popular Python AI Tech Stacks.

DEVELOP INTELLIGENT SEARCH & CHAT WITH VECTOR SEARCH

Leverage Vector Search and the popular LangChain+HuggingFace stack. Install PyMongo first.

Connect to Cluster0

Set up connection security Choose a connection method Connect

Connecting with MongoDB Compass

I don't have MongoDB Compass installed I have MongoDB Compass installed

1. Choose your version of Compass

1.12 or later

See your Compass version in "About Compass"

2. Copy the connection string, then open MongoDB Compass

`mongodb+srv://cheerio:<password>@Cluster0.so7kree.mongodb.net/`

Replace <password> with the password for the **cheerio** user. Ensure any options are URL encoded.

RESOURCES

Connect with Compass Import and Export Data

Access your Database Users Troubleshoot Connections

Go Back Close Review setup steps

Featured Resources

PyMongo

Aggregations in Python

Data Encryption Quickstart

PyMongoArrow: Data Analysis

Beanie ODM Quickstart

More Python Content

Sample Apps

PyMongo

Flask Stack

FAST API

Pydantic Stack

New On Atlas

2 NEW

Learn about the latest feature enhancements on Atlas.

MongoDB Compass Connection

The screenshot shows the MongoDB Compass application running within a Visual Studio Code window. The interface is divided into several sections:

- Left Sidebar (EXPLORER):** Displays the file structure of the project, including folders like 'GROUP 1', 'GROUP 2', and 'BIGDATA', and files like 'main.py', 'README.md', 'chart_description', 'cut_data.bash', 'requirements', 'datasets', 'env', 'gq_2010_us_040', 'image_tem.odp', 'main.py', 'playground-1.m', 'README.md', 'requirements.tx', 'cidc_calc', 'cidc-multiple-run', 'flask_demo', 'FP_temp.depreca', 'hw3_json_app_re', '__pycache__', 'env_setup', 'planning', 'OUTLINE', and 'TIMELINE'.
- Top Bar:** Shows the title 'main.py - BigData - Visual Studio Code' and standard window controls.
- MongoDB Compass Window:**
 - Connect Tab:** Features a 'New connection +' button and a 'Saved connections' list with 'PrecisionPC' (Apr 28, 2024, 7:54 PM) selected.
 - Main Panel:**
 - Header:** 'PrecisionPC' with a 'FAVORITE' star icon.
 - URI:** A text field containing 'mongodb+srv://cheerio:****@cluster0.so7kree.mongodb.net/'.
 - Advanced Connection Options:** A tabbed interface with 'General' selected. It shows 'Connection String Scheme' with 'mongodb' and 'mongodb+srv' options, and 'Hostname' with 'cluster0.so7kree.mongodb.net'.
 - Buttons:** 'Save' and 'Connect' buttons at the bottom right.
 - Right Panel:** Contains informational text: 'New to Compass and don't have a cluster?', 'How do I find my connection string in Atlas?', and 'How do I format my connection string?'.
- Bottom Panel:** Shows the terminal output, including the command 'cheerio@cheerio-precision:~/Rowan_Local/BigData/BigDataFP (main)\$'.

MongoDB Charts

The screenshot displays the MongoDB Atlas web interface. At the top, the navigation bar includes the Atlas logo, a dropdown menu for 'Peyton's Or...', and links for 'Access Manager' and 'Billing'. On the right side of the top bar are links for 'All Clusters', 'Get Help', and a user profile dropdown for 'Peyton'.

Below the top bar, a secondary navigation bar shows 'Project 0' and tabs for 'Data Services', 'App Services', and 'Charts' (which is currently selected). On the far right of this bar are icons for user management, sharing, and notifications.

The main content area is titled 'Dashboards'. On the left is a sidebar with a tree view containing categories: 'DASHBOARDS' (with sub-items 'Project', 'Organization', and 'Owners'), 'DEPLOYMENTS' (with 'Data sources' and 'Ingestions'), 'DEVELOPMENT' (with 'Embedding'), 'SCHEDULING' (with 'Reports'), 'SETTINGS' (with 'AI features'), and 'USAGE' (with 'Data transfer').

The 'Dashboards' section includes a filter bar with 'All dashboards', 'Last modified date', a search input 'Title or description', and an 'Add Dashboard' button. A dashboard card titled 'Peyton's Dashboard' is shown, featuring a lock icon and a menu icon. It lists two charts: 'Comparative Analysis of Average Air Quality Index ...' and 'air_quality', followed by 'AQI By Coord in 2017'. At the bottom of the card, it indicates '6 charts' and 'Modified 2 days ago'.

MongoDB Charts

The screenshot displays the MongoDB Atlas user interface. At the top, the navigation bar includes the Atlas logo, a dropdown menu for 'Peyton's Or...', and links for 'Access Manager' and 'Billing'. On the right side of the navigation bar, there are links for 'All Clusters', 'Get Help', and a user profile dropdown for 'Peyton'. Below the navigation bar, a secondary menu shows 'Project 0', 'Data Services', 'App Services', and 'Charts' (which is highlighted). The main content area is titled 'Peyton's Dashboard' and features a large chart placeholder. The chart is titled 'Comparative Analysis of Average Air Quality Index Across States' with a subtitle '2017-2021'. A 'Loading chart...' message is centered in the placeholder. To the right of the chart, there is a toolbar with icons for 'More', 'Refresh', 'Fullscreen', 'Y-axis', 'Schedule', 'Share', and 'Add Chart'. The 'Add Chart' button is highlighted with a green circle.

Atlas Peyton's Or... Access Manager Billing All Clusters Get Help Peyton

Project 0 Data Services App Services Charts

← Peyton's Dashboard ... Refresh Fullscreen Y-axis Schedule Share Add Chart

Comparative Analysis of Average Air Quality Index Across States
2017-2021

Loading chart...

MongoDB Atlas Charts

The screenshot displays the 'Add Chart' interface in MongoDB Atlas. The interface is divided into several sections: a top navigation bar with 'Classic' and 'Natural Language' tabs, and a 'PREVIEW' button; a 'Data Source' section with a 'Select' button; a 'Query' section with a 'Library' dropdown and a text input field; a 'FIELDS' section with a search bar; a 'CHART TYPE' section with a 'Grouped Bar' dropdown; and a 'PREVIEW' section with 'COMPLETE' and 'SUBSET' tabs. A 'Select Data Source' dialog box is open in the center, showing a 'Suggested' tab with a list of data sources used in the dashboard, including 'air_quality.air_quality'. The dialog box has 'Select' and 'Cancel' buttons.

Add Chart
Peyton's Dashboard

Classic Natural Language PREVIEW

Get help Cancel Save and close

Data Source: Select

Query: Library Type a (query) or an [aggregation] to pre-process your data Apply

FIELDS: Name

CHART TYPE: Grouped Bar

PREVIEW COMPLETE SUBSET

Select Data Source

Suggested Project Sample

Data sources used in this dashboard

air_quality.air_quality Select

Cancel

MongoDB Atlas Charts

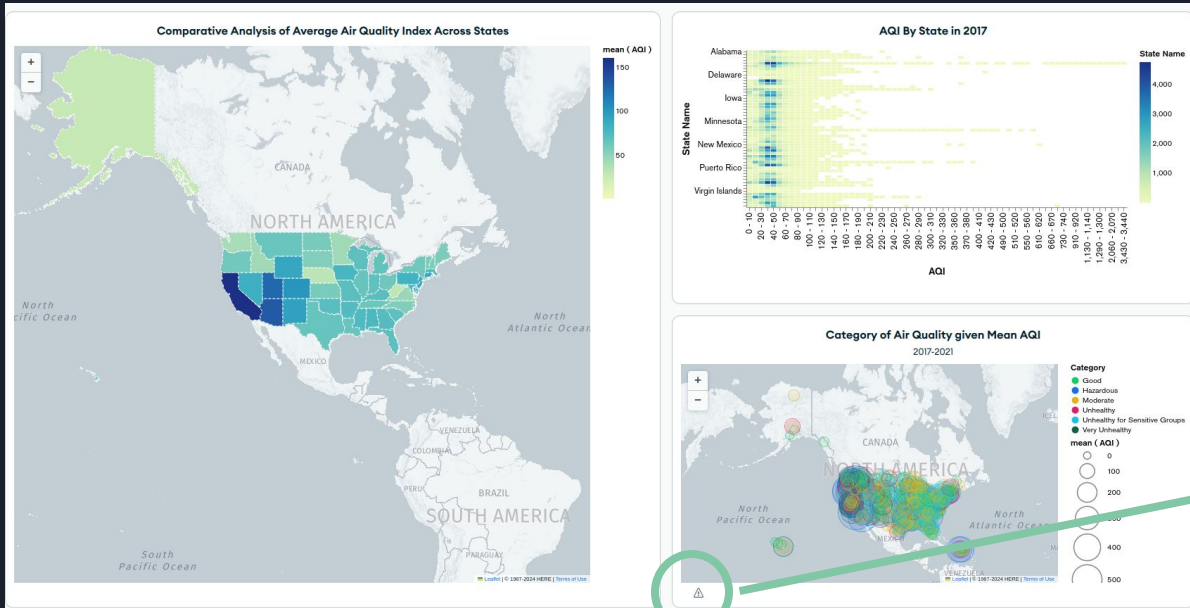
The screenshot displays the 'Add Chart' interface in MongoDB Atlas. The 'Data Source' is set to 'air_quality'. The 'Query' field is empty, with a placeholder text: 'Type a {query} or an [aggregation] to pre-process your data'. The 'CHART TYPE' dropdown is set to 'Grouped Bar', which is circled in green. A green arrow points from this dropdown to a callout window that shows the full list of available chart types. The callout window lists the following categories and options:

- CHART TYPE**
 - Grouped Bar (selected)
 - Stacked Combo
- LINE**
 - Discrete Line
 - Continuous Line
- AREA**
 - Discrete Area (highlighted)
 - 100% Stacked Area
 - Continuous Area
- GRID**
 - Heatmap
- SCATTER**
 - Scatter
- CIRCULAR**
 - Donut

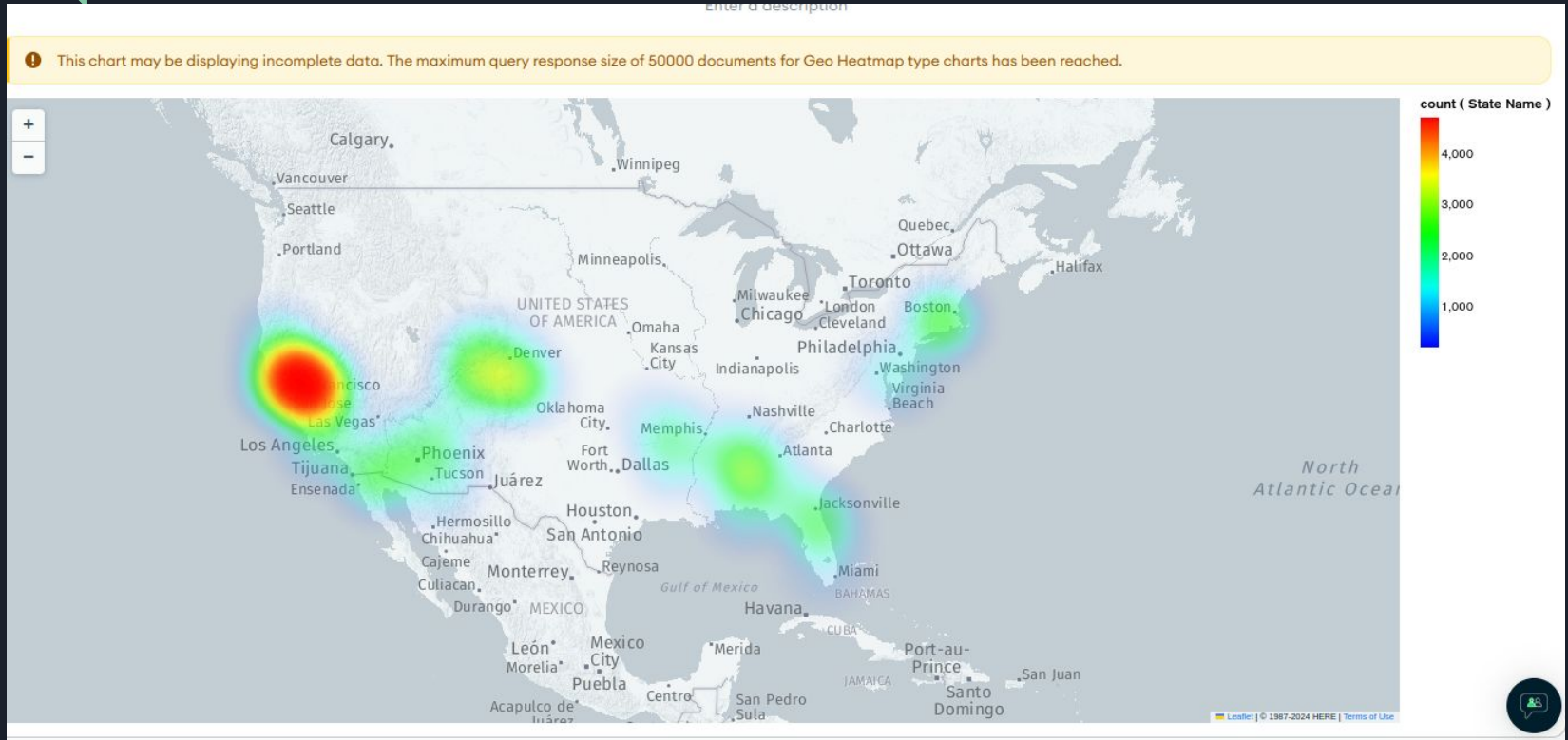
The main interface also shows a 'FIELDS' list on the left, including '_id', 'AQI', 'Category', 'County Name', 'Date', 'Defining Parameter', 'Latitude', 'Longitude', and 'State Name'. The 'CHART TYPE' dropdown is set to 'Grouped Bar'. The 'Preview' section shows a large number '1,000' and a small bar chart. The 'APPLY' button is visible at the bottom right of the preview section.

MongoDB Chart Limitations

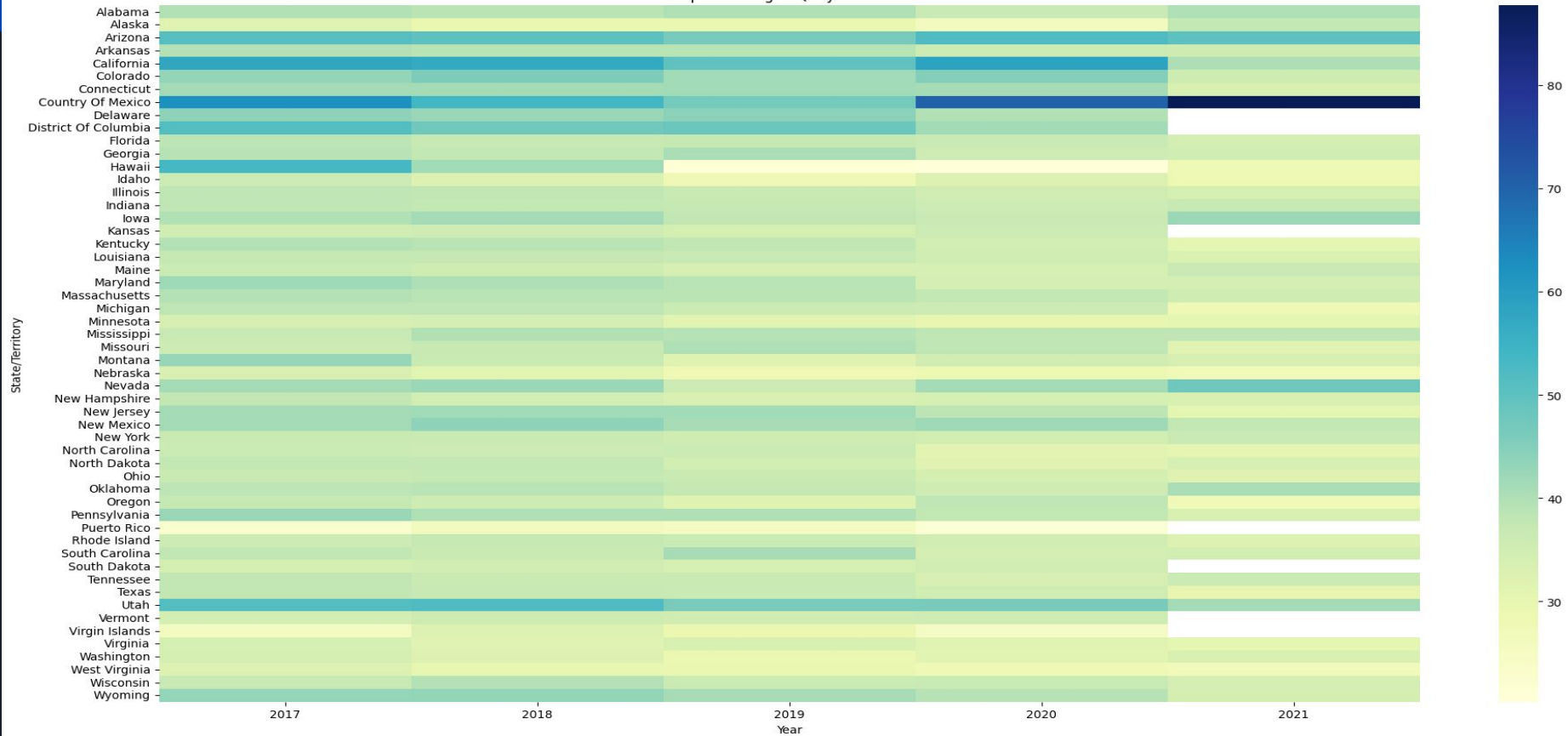
! This chart may be displaying incomplete data. The maximum query response size of 1000 documents for Geo Scatter type charts has been reached.



MongoDB Chart Limitations



Heatmap of Average AQI by State and Year



```
new_data.predictions = model.predict(new_data)
predicted_categories = le.inverse_transform(new_data.predictions)
print(f"Predicted categories for new data: {predicted_categories}")
```

Model Accuracy: 1.0

Predicted categories for new data: ['Good' 'Unhealthy for Sensitive Groups' 'Very Unhealthy']

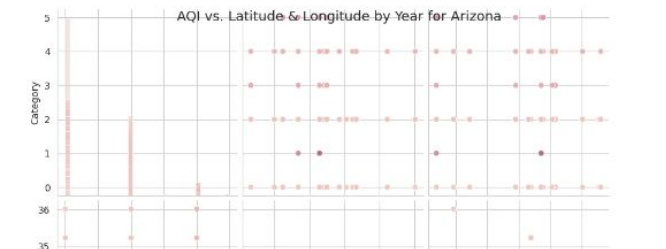
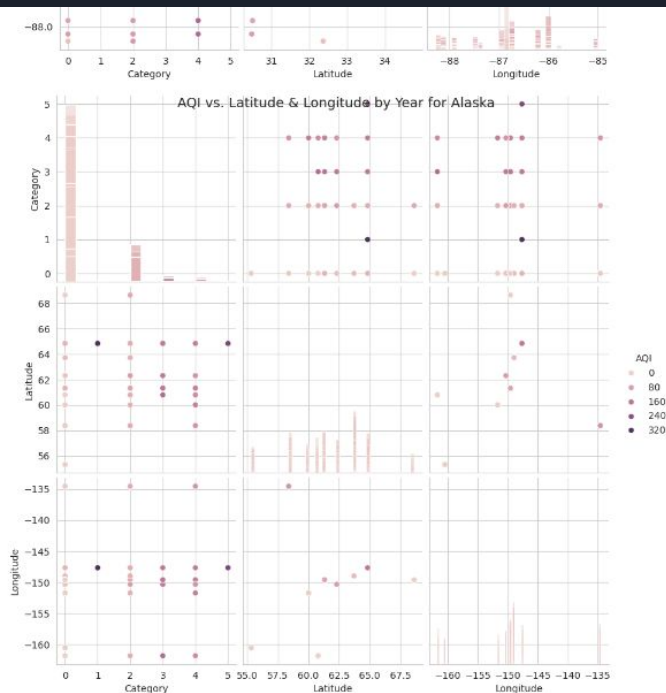
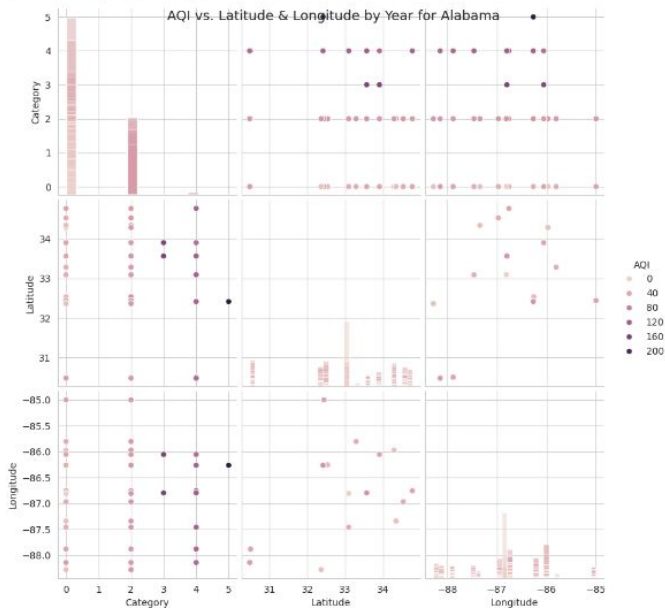
```
In [25]: # Function to create a pairplot for a specific state
def state_pairplot(state_data):
    # Filter data for the specific state
    data_filtered = state_data[state_data['State Name'] == state]

    # Create a pairplot with AQI as hue variable (colors by AQI value)
    sns.pairplot(data_filtered, hue='AQI', height=3, diag_kind='hist') # Updated cell

    # Add a title with state name
    plt.suptitle(f"AQI vs. Latitude & Longitude by Year for {state}", fontsize=14)
    plt.show()

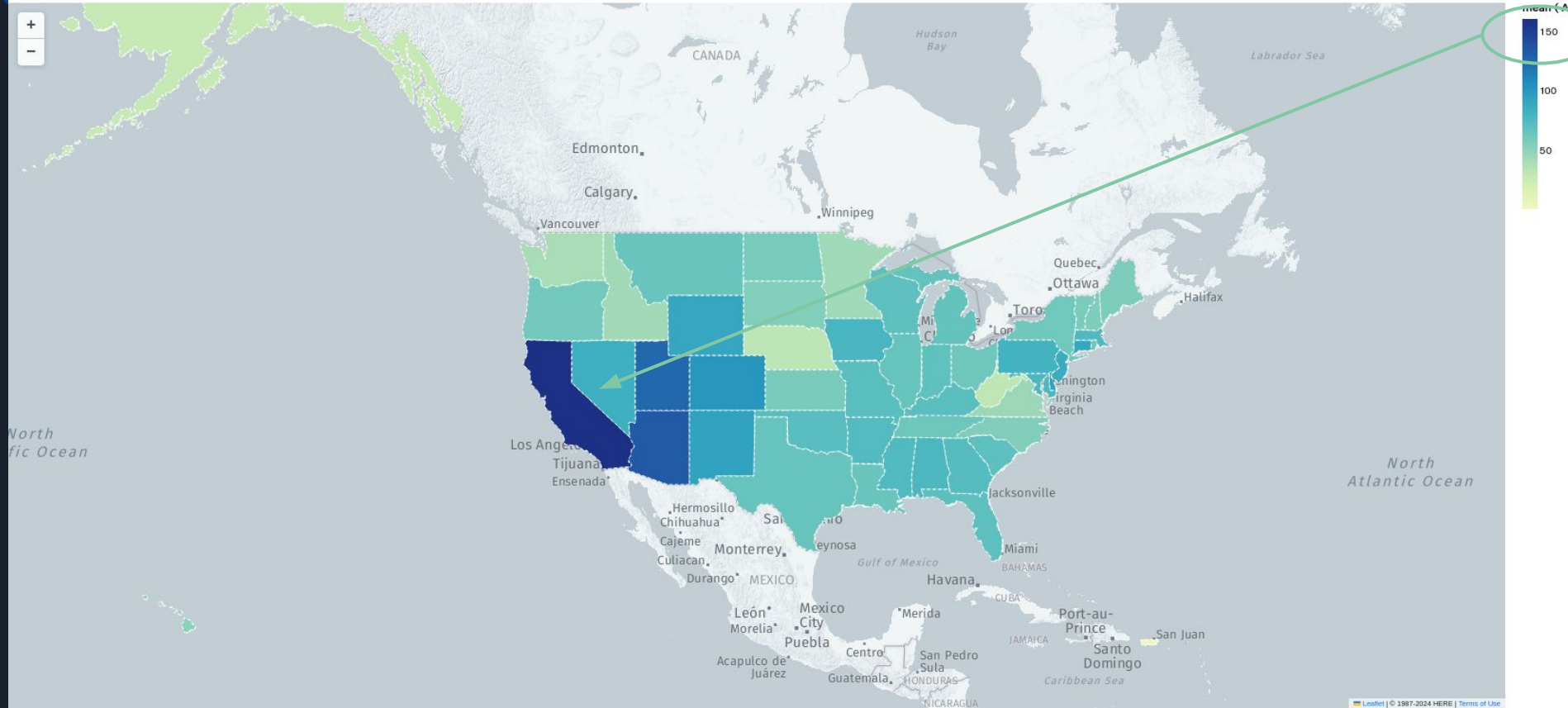
    # Group data by state (assuming 'State Name' is the grouping variable)
    state_groups = data.groupby('State Name')

    # Iterate through each state group
    for state, state_data in state_groups:
        state_pairplot(state_data.copy()) # Pass a copy to avoid modifying original data
```



Comparative Analysis of Average Air Quality Index Across States

2017-2021



2017-2021

CHART TYPE

 Geo Choropleth

Encode

Filter 1

Customize

Location

⋮ **A** State Name

☐ Binning

Color

⋮ # AQI

Aggregate

MEAN

Encode

Filter 1

Customize

General

COLOR PALETTE

☐ Custom Color Palette

CHART ELEMENTS

☒ Automatically Set Viewport

Label Size

100%

Opacity

100%

Shape Scheme

US States

Canadian Provinces

Dutch Provinces

German States

mean (A

150

100

50

AQI Variation over Time for each Territory/State

1/2017 - 12/2020

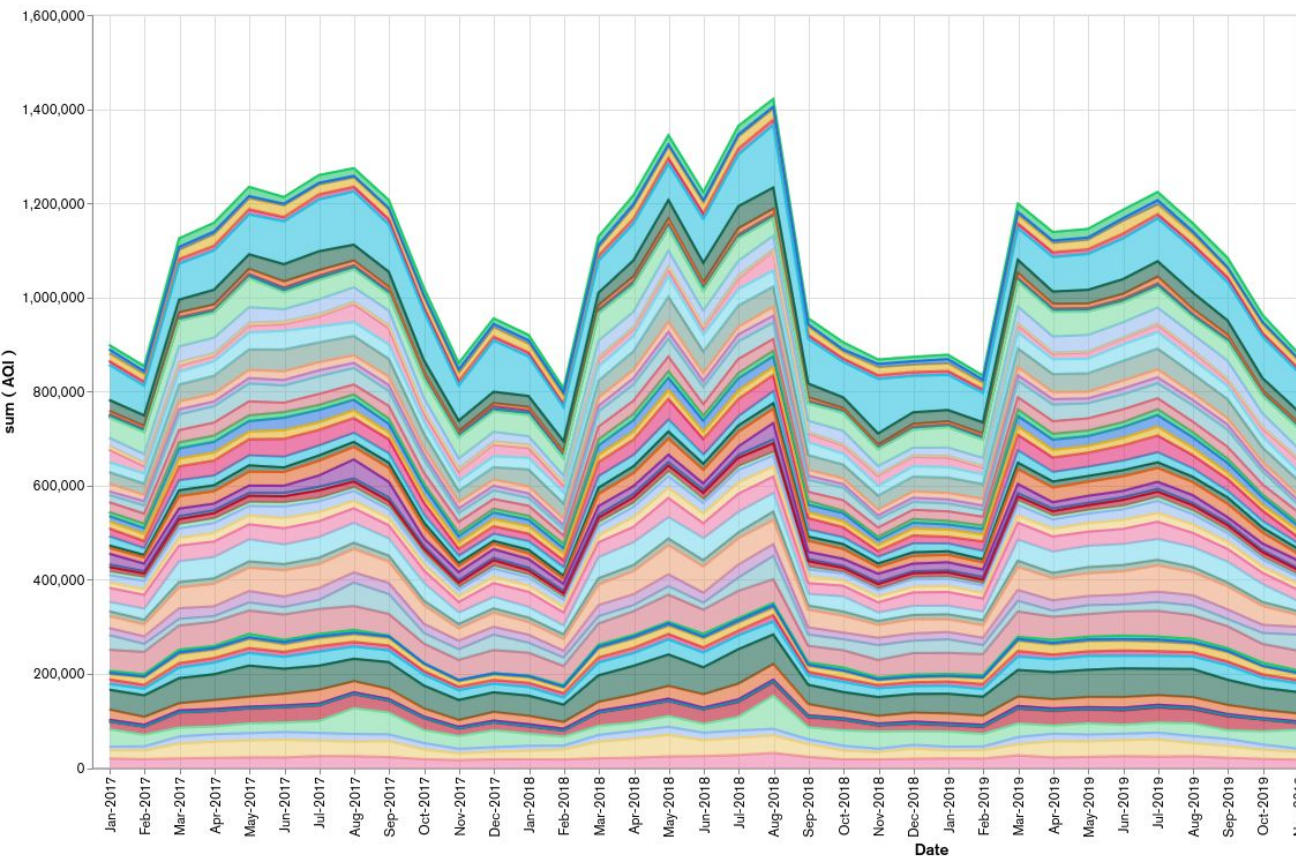


CHART TYPE

Discrete Area

Encode
Filter 1
Customize

X Axis

Date

☒ Binning

MONTH

☐ Periodic

Y Axis

AQI

Aggregate

SUM

Series

A State Name

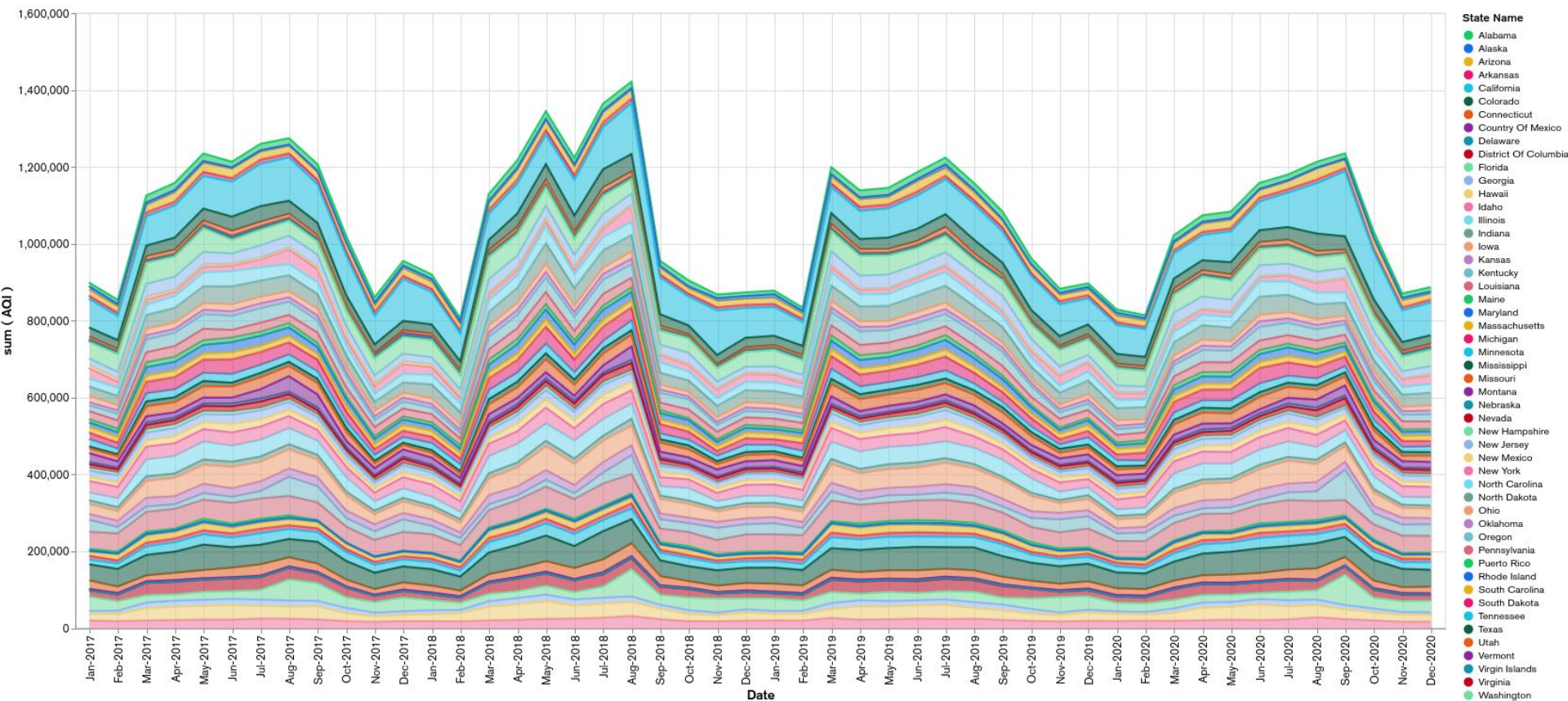
☐ Binning

State Name

- Alabama
- Alaska
- Arizona
- Arkansas
- California
- Colorado
- Connecticut
- Country Of Mexico
- Delaware
- District Of Columbia
- Florida
- Georgia
- Hawaii
- Idaho
- Illinois
- Indiana
- Iowa
- Kansas
- Kentucky
- Louisiana
- Maine
- Maryland
- Massachusetts
- Michigan
- Minnesota
- Mississippi
- Missouri
- Montana
- Nebraska
- Nevada
- New Hampshire
- New Jersey
- New Mexico
- New York
- North Carolina
- North Dakota
- Ohio
- Oklahoma
- Oregon
- Pennsylvania
- Puerto Rico
- Rhode Island
- South Carolina
- South Dakota
- Tennessee
- Texas
- Utah
- Vermont
- Virgin Islands
- Virginia
- Washington

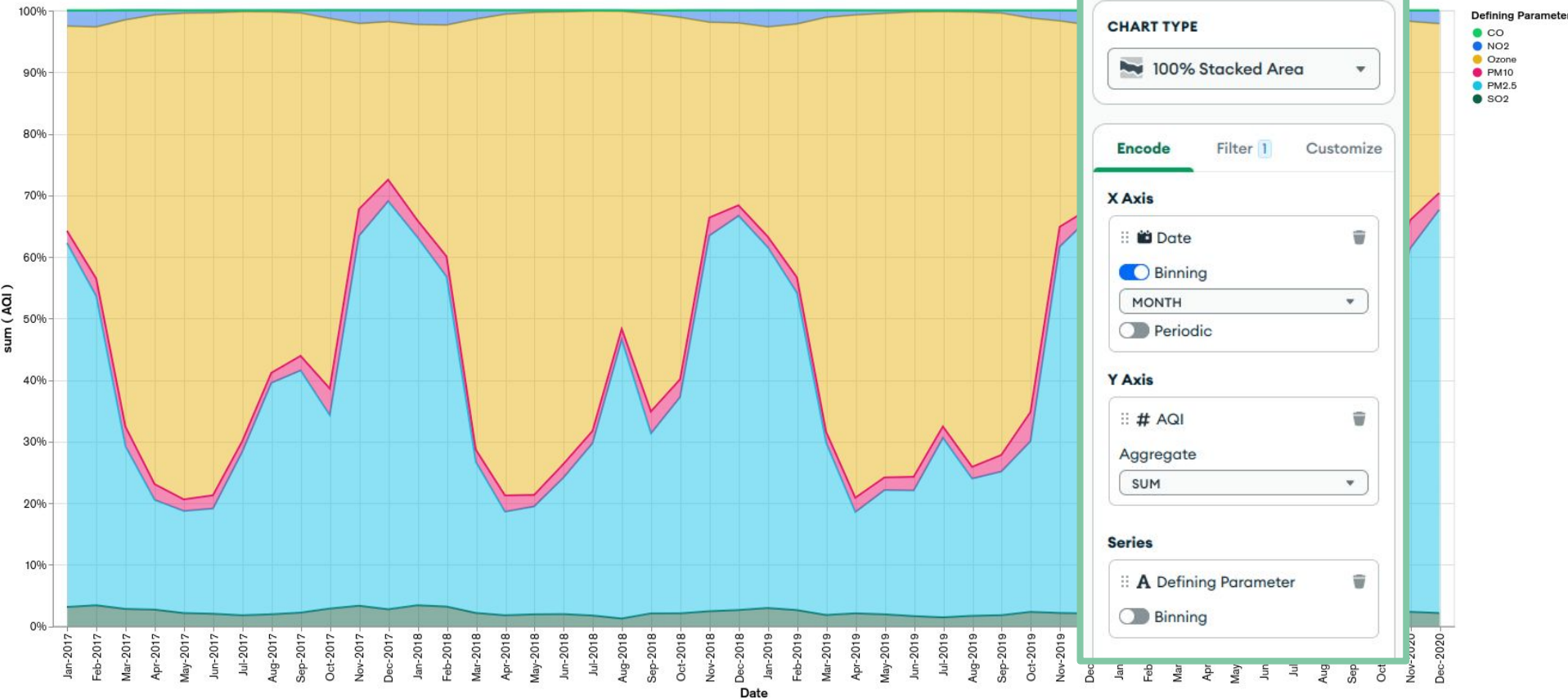
AQI Variation over Time for each Territory/State

1/2017 - 12/2020



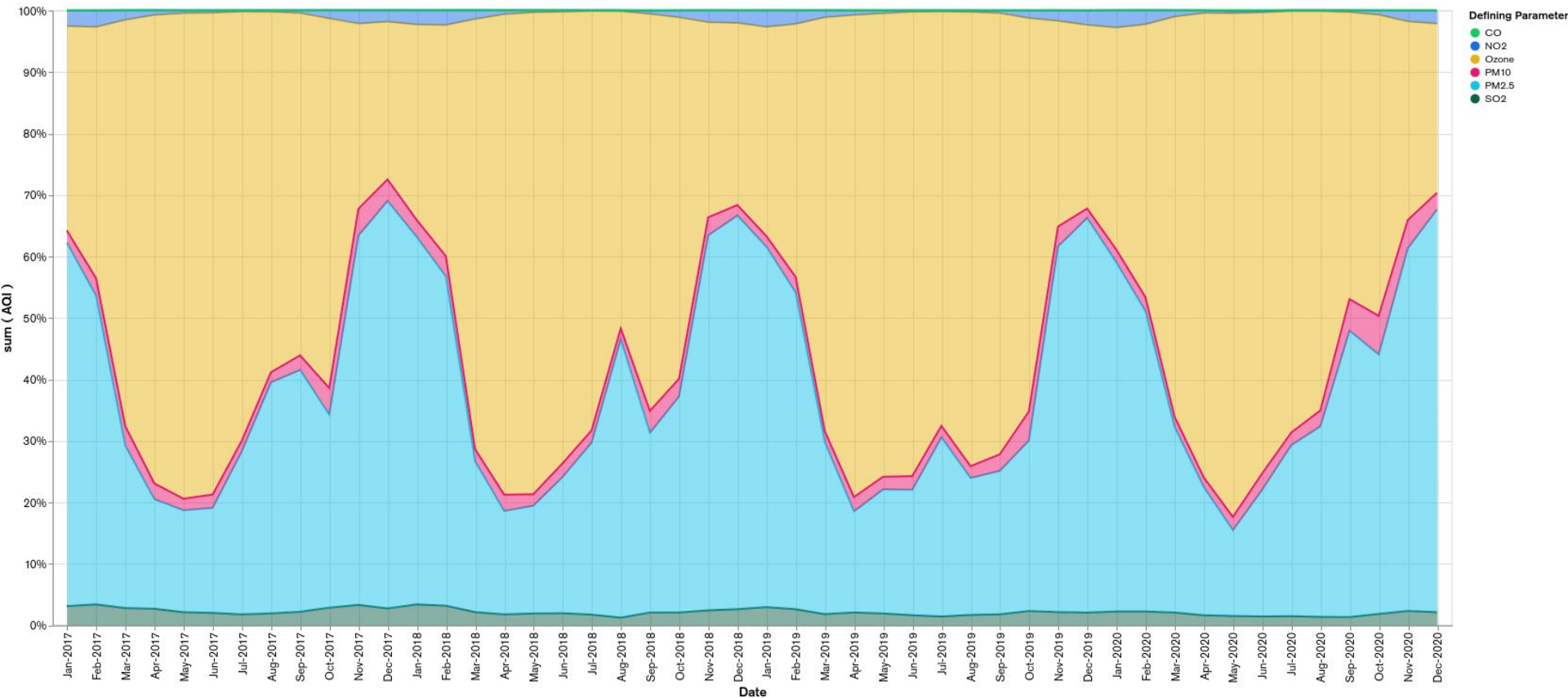
Seasonal and Annual Percentage Changes in Defining Parameter

2017-2020



Seasonal and Annual Percentage Changes in Defining Parameter

2017-2020





Github Repository

<https://github.com/inkieee/BigDataFP>



Like/Dislike

- Liked
 - Exposure to a lot of different tools
 - CICD
 - Distributed HA concepts
 - Various DB types
 - Example git repos
- Disliked
 - Required using up free trials to complete assignments
 - Couldn't download gitbook easily



Thanks!



Following slides here for posterity



GitHub Limitations

```
size of 50.00 MB
remote: error: Trace: 45002baf118dcfb1debbb79f1ed4104a1cf17e4fadb257724ef3eb4298a3e8ce
remote: error: See https://gh.io/lfs for more information.
remote: error: File data/air_quality/air_quality/aqi_daily_2011_to_2021.csv is 181.84 MB; this exceeds GitHub's file size limit of 100.0
) MB
remote: error: File data/air_quality/air_quality/aqi_daily_1980_to_2021.csv is 638.04 MB; this exceeds GitHub's file size limit of 100.0
) MB
```



GitHub Limitations

“Additionally, had some issues with uploading document data into github for free for too large file sizes.”

Grafana Attempt

```
(Reading database ... 215041 files and directories currently installed.)
Preparing to unpack grafana_10.4.2_amd64.deb ...
Unpacking grafana (10.4.2) ...
Setting up grafana (10.4.2) ...
Adding system user `grafana' (UID 135) ...
Adding new user `grafana' (UID 135) with group `grafana' ...
Not creating home directory `/usr/share/grafana'.
### NOT starting on installation, please execute the following statements to con
figure grafana to start automatically using systemd
  sudo /bin/systemctl daemon-reload
  sudo /bin/systemctl enable grafana-server
### You can start grafana-server by executing
  sudo /bin/systemctl start grafana-server
```

Installing on a local Grafana:

For local instances, plugins are installed and updated via a simple CLI command. Plugins are not updated automatically, however you will be notified when updates are available right within your Grafana.


1. Install the Data Source

Use the grafana-cli tool to install MongoDB from the commandline:

```
grafana-cli plugins install grafana-mongodb-datasource
```

Enterprise
only

Mongo w/ Grafana

 You have uncapped usage until May 12, 2024. Upgrade plans to continue using Grafana Cloud with unlimited, pay-as-you-go usage. [Upgrade now](#)

Looks exactly like
free self hosted :v

Plugins

Extend the Grafana experience with panel plugins and apps. To find more data sources go to [Connections](#).

Search

Q mongo [x Clear](#)



MongoDB





By Grafana Labs

Enterprise

grafana-mongodb-datasource-2

Type: MongoDB


[Settings](#) [Permissions](#) [Insights](#) [Cache](#)


Name  grafana-mongodb-datasource-2 [Default](#) 

Before you can use the MongoDB data source, you must configure it below or in the config file. For detailed instructions, [view the documentation](#).

Fields marked with * are required

Connection

Connection string *  mongodb+srv://cluster.host.net/dbname?retryWrites+1


 Please enter a valid URL


Authentication

Authentication methods

Choose an authentication method to access the data source

Credentials

User *  User

Password *  Password



Grafana Attempt

“

Initial project attempts included Grafana and Tableau, all of which were new tools to me, which I encountered other issues. For example Grafana requires a mongodb-plugin that is only available to Enterprise accounts of grafana, even when hosted locally from your machine, the plugin verifies against some enterprise key and if incorrect, will not allow mongodb connections. You can use the online Grafana SaaS service which also has it's own payment structure like Azure or MongoAtlas, but also requires an enterprise key. You can use a free trial for 2 weeks though.

”

Mongo w/ Grafana

Overview

DEPLOYMENT

Database

Data Lake

SERVICES

Device Sync

Triggers

Data API

Data Federation

Atlas Search

Stream Processing

Migration

SECURITY

Backup

Database Access

Network Access

Advanced

PEYTON'S ORG - 2023-10-21 > PROJECT 0

Network Access

IP Access ListPeeringPrivate Endpoint

+ ADD IP ADDRESS

You will only be able to connect to your cluster from the following list of IP Addresses:

IP Address	Comment	Status	Actions
<input type="text"/>	My IP Address	● Active	<button>⚙ EDIT</button> <button>🗑 DELETE</button>
0.0.0.0/0 (includes your current IP address)		● Active	<button>⚙ EDIT</button> <button>🗑 DELETE</button>



Mongo w/ Grafana

“

Being hosted on Grafana servers though, we do have to modify our MongoDB network connection profile to be open to the internet (0.0.0.0 netmask 0) which allows all connections (including from grafana IPs).”

Grafana issues w/ Cloud

The screenshot displays the Grafana web interface in a dark theme. The top navigation bar includes a search bar with the placeholder text "Search or jump to..." and a "ctrl+k" shortcut. Below the navigation bar, the breadcrumb trail reads "Home > Dashboards > New dashboard > Edit panel". The main panel area shows a large empty space with the text "No data". The left sidebar contains a navigation menu with categories like "Home", "Starred", "Dashboards", "Reporting", "Explore", "Alerts & IRL", "OnCall", "Incidents", "Machine learning", "SLO", "Testing & synthetics", "Infrastructure", "Kubernetes", "AWS", and "Application". The right sidebar contains a "Time series" panel with a search bar and various options like "Panel options", "Panel links", "Repeat options", "Tooltip", and "Legend". The bottom panel shows a query editor with a "Data source" dropdown menu open, listing several data sources: "grafana-mongodb-datasource", "grafana-mongodb-datasource-1", "grafanacloud-cheerio-alert-state-history", and "grafanacloud-cheerio-cardinality-management". The query editor also includes a "Query options" section with "MD = auto = 1187" and "Interval = 20s", and a "Query inspector" button. A "Run query" button is visible at the bottom right of the query editor.

Home > Dashboards > New dashboard > Edit panel

Table view Fill Actual Last 6 hours Time series

Panel Title

No data

Query 1 Transform data 0 Alert 0

Data source grafana-mongodb-datasource Query options MD = auto = 1187 Interval = 20s Query inspector

grafana-mongodb-datasource MongoDB

grafana-mongodb-datasource-1 MongoDB

grafanacloud-cheerio-alert-state-history Loki

grafanacloud-cheerio-cardinality-management Grafana cardinality

Run query

Open advanced data source picker

bad query (1773444444) (Trace ID: sh3a31d56a9df19a7d5a9e10a42f2a5)

Search options All Overrides

Panel options

Title Panel Title

Description

Transparent background

Panel links

Repeat options

Tooltip

Tooltip mode Single All Hidden

Hover proximity

How close the cursor must be to a point to trigger the tooltip, in pixels

Max height

Legend

Visibility



Grafana issues w/ Cloud

“

Annoyingly I had issues with generating queries of any type and decided not to invest too much time on something that would expire in 2 weeks anyway.

“