Analysis Of Storms

Team:

Hassan Aftab
Lindsey Krempa
Mason DeJesus
Eugenio Elizondo
Thripura Pakala

Introduction:

This project is to demonstrate data cleaning, data analysis and data presentation skills to show the functional data application that tells an interactive compelling data story and building a full stack application. A flask backend that can query against a SQLITE database using user inputs/filter. A HTML/JS front-end to facilitate user requests and to dynamically build visualization and creating an interactive dashboard to drive Persianized discoveries of the dataset and including a map to showcase GIS and GEO visualization techniques, which team grasped the concepts provided the past 3 months from the Data analysis and Visualization Certification Course. We chose the Storms dataset because we found a common interest in Storms and how they are impacting the world based on Global Warming and how the storms are strengthening over the period of time1852-2021, how it is impacting the weather we live in. This presentation aimed to highlight the significance of creating Home, Dashboard, Map, About Us and Work cited html pages, data collection and effective data visualization for informing decision-making and driving insights.

Project Scope:

The Scope was limited to analysis on a CSV dataset; no more than 10MB of data. A minimum of three research questions were answered including inspiration as we built upon previous topics. The dataset chosen came from Kaggle.com as Kaggle has a wealth of data and previous analysis to build upon. Three visualizations and a map were included yet will be considered. The visualizations included Stacked Bar chart, line graphs, map and other useful visualizations to accommodate proper presentation of the distribution, spread, of the data. Roles and responsibilities were initially defined and presented in the project proposal with modifications as needed throughout the project.

Color Palette:

A color pallet was included for consistency across the graphs.

https://www.color-hex.com/color-palette/114111

Dataset Cleaning (ETL):

Upon initial analysis of the Storms dataset, we begin by inspecting the dataset to understand its structure, variables and any potential issues. Storms dataset which we got from Kaggle, Kaggle dataset accuracy and integrity is very good, we have Storms data from 1852-2021. There are no null values, no duplicates, no inconsistencies in the data entry, no typos, no misspellings in the dataset. Our team decided to drop two columns that are not very useful in the analysis so dropped these tropicalstorm_force_diameter, hurricane_force_diameter columns. Storms data has Year, Month, Day given separately, our team decided to join them together to make a date column using this code:

storms_df['combined_date'] = storms_df. apply (**lambda** x: pd.to_datetime(f"{x['year']}-{x['month']}-{x['day']}", format='%Y-%m-%d'), axis=1)

We arranged the columns as we wanted with this code:

new_columns = ['name', 'combined_date', 'year', 'month', 'day', 'hour', 'lat', 'long', 'status', 'category', 'wind', 'pressure']

storms_df = storms_df[new_columns]

After ETL is done converted the file into csv using this code:

storms_df.to_csv ("Resources/storms_final.csv", encoding='utf8', index=**False**)

HTML Pages:

This is our website link:

https://56f7-2603-8081-4e06-7f4-e9c7-561e-395-47b2.ngrok-free.app/

Data Flow:

Data Cleaning in Jupyter Notebooks: Data is cleaned and prepared in Jupyter Notebooks before being stored in the SQLite database.

SQLAlchemy: Handles the ORM (Object-Relational Mapping) between Python classes and database tables, making it easier to work with database data in the Flask application.

SQLite Database: Stores the cleaned data that will be displayed on the website.

Flask Application: Retrieves data from the SQLite database using SQLAlchemy and sends it to the front end for display.

Frontend Components:

HTML Files: These files will contain the structure and layout of the web pages.

CSS Files: Used for styling the HTML elements.

JavaScript Files: App.js used for adding interactivity to the website, such as fetching data from APIs or handling user interactions.

Backend Components:

Flask Application: App.py contains the Flask application code, handles routing, request processing, responses and logic for the Web application.

SQLAlchemy: Used for interacting with the SQLite database and mapping database tables to Python classes.

SQLite Database: Stores the data that needs to be accessed by the website.

SQLiteHELPER.py: Contains functions to interact with the SQLite databases such as querying data or updating records.

Interaction Flow:

When a user accesses the website, the Flask application processes the request.

The Flask application interacts with the SQLite database using SQLAlchemy to retrieve or update data.

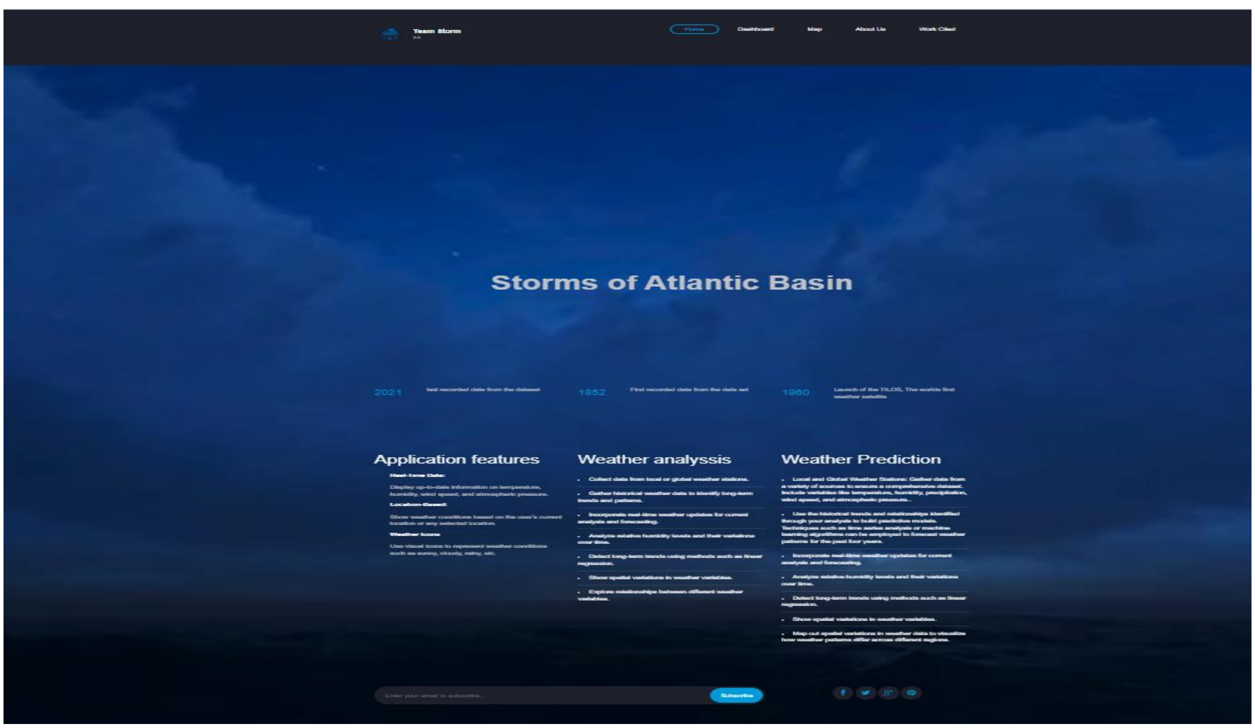
The Flask application renders the HTML templates based on the retrieved data and sends the response back to the user's browser.

Website Architecture Diagram:

A visual representation of the components and how they interact with each other can be created to illustrate the website architecture. By integrating these components and defining the flow of data and interactions, we created a comprehensive website architecture for our project.

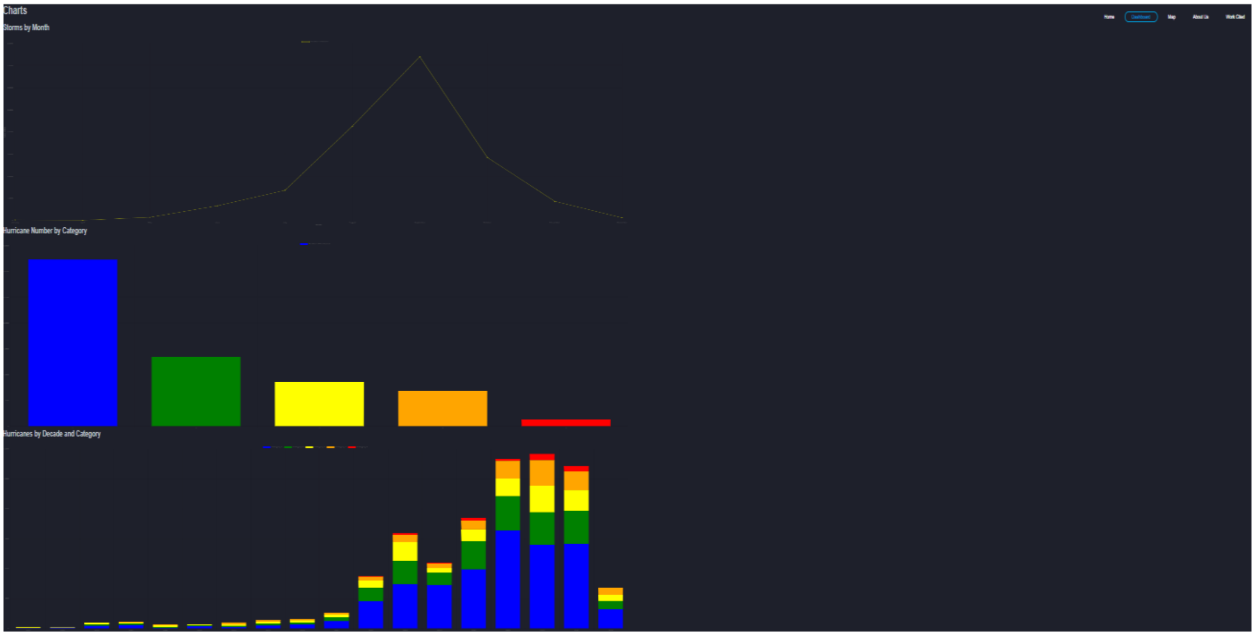
Home Page:

This is our main home page. Here, we have links to the Dashboard page, Map, About Us, and Work Cited pages. We have added some live floating clouds and thunder effects to the background of our home page. Users can navigate from the main page to any other page to view additional information. There is a heading "Storms of the Atlantic Basin" with details about storms from 1852, 1960, and 2021, explaining features like application features, weather analysis, and weather prediction. Additionally, there is a subscribe button and links to Facebook, Twitter, Google, and Pinterest.



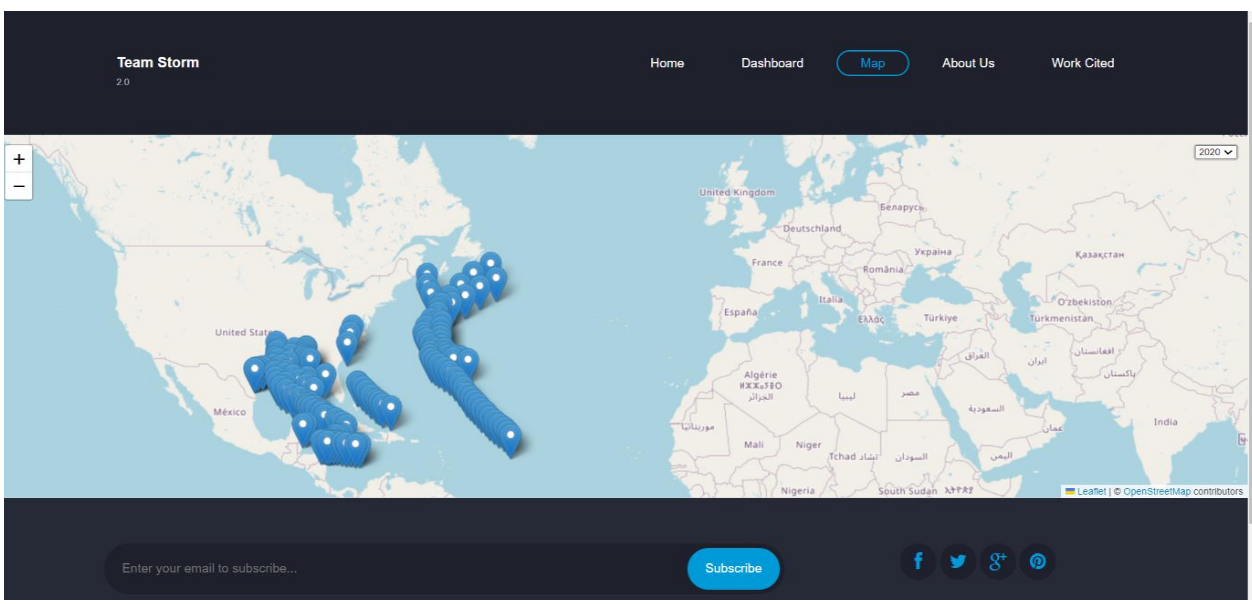
Dashboard Page:

This is our Dashboard page, which contains information about Charts that answer our research questions. Detailed information is provided with each chart below.



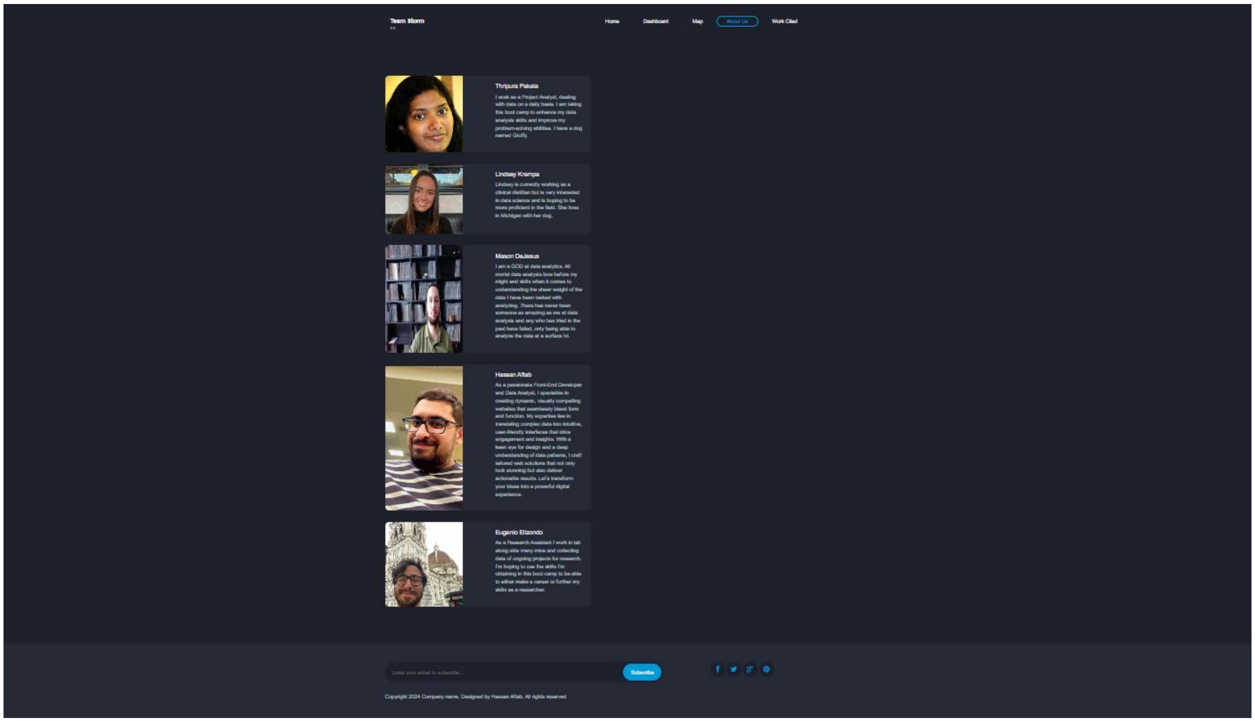
MAP:

This is our Map that shows the storms that occurred in the Atlantic Basin. We have added filters to the chart to view the number of storms that occurred in a specific year. (1950-2021)



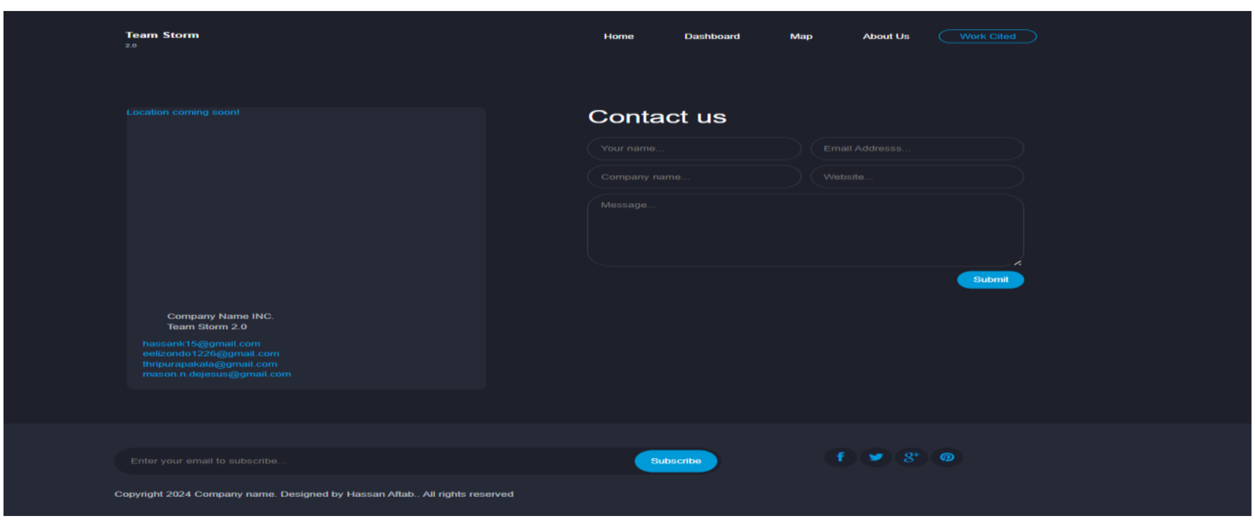
About Us Page:

The About Us page will provide information about our team who worked on this project and a little bit of information about what we do.



Contact Us Page:

We have a Contact Us page that we didn't have enough time to update.



Visualizations:

Through data visualizations we show what potentially increases chances of Storms and answers data questions through Data Visualization.

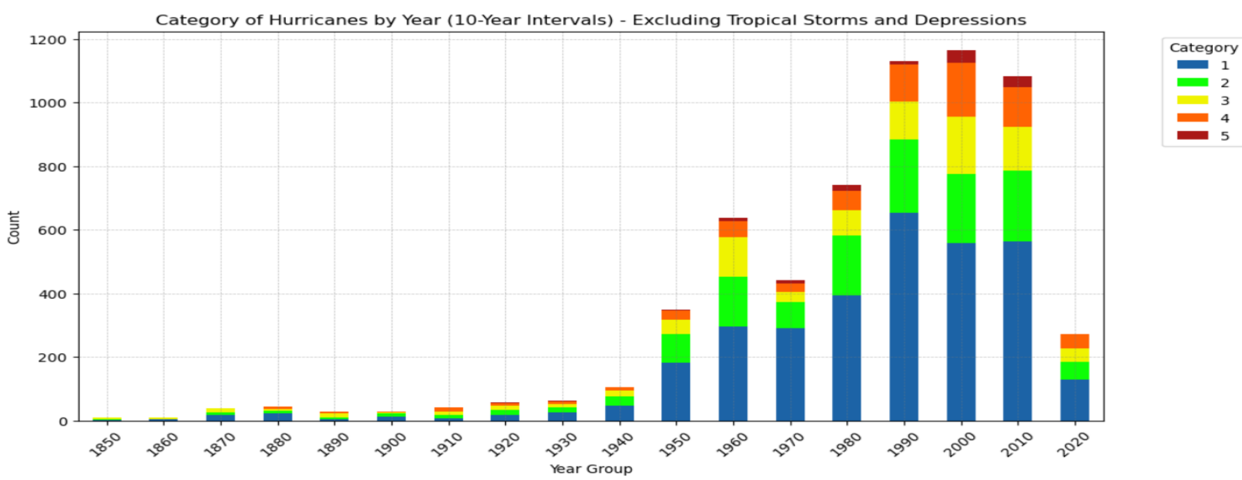
Below are questions which will be answered through Visualization:

1. Has there been an increase or a decrease in the frequency of storms over time?
2. Are there any trends in the intensity of storms impacting the Atlantic Basin?
3. Are there any trends in the appearance or movement of Hurricanes vs Categories?
4. Which seasons affect the Storms over the period (1852-2021)?

1. Has there been an increase or a decrease in the frequency of storms over time?

As per the data the No of Storms with 10-year Intervals vs No of Categories shows storms are increasing over the period.

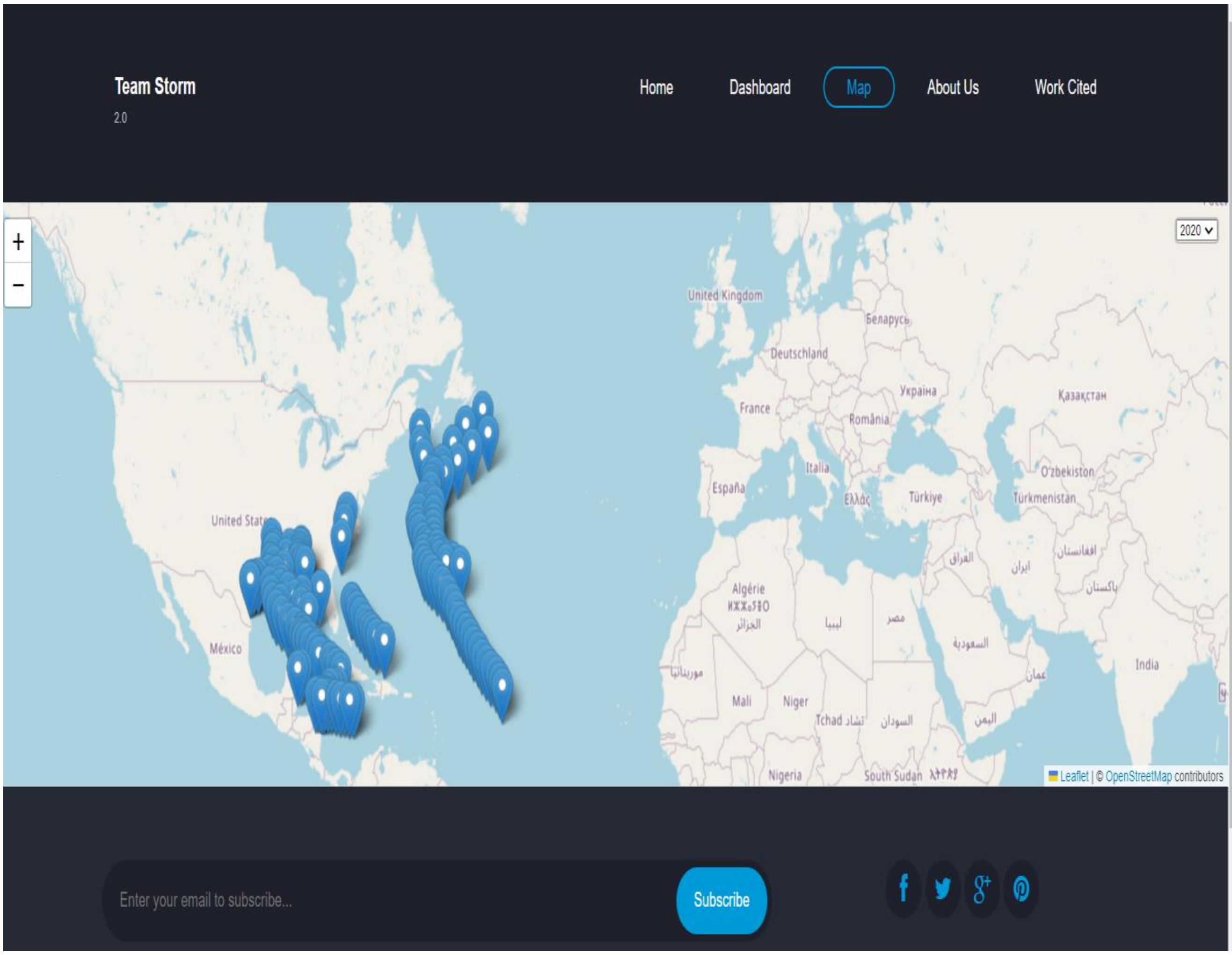
- Introduction: Analysis of storm frequency over time reveals interesting trends across different storm categories.
- Category 1: Category 1 storms peaked in frequency in 1990 but experienced a significant decrease from 2010 to 2020.
- Category 2: Category 2 storms maintained a consistent frequency from 1980 to 2010.
- Category 3 & 4: Category 3 and 4 storms showed a steady occurrence from 1970 to 2020.
- Category 5: Category 5 storms exhibited an increasing trend until 2010, followed by a notable drop in frequency in 2020.
- The insights from the chart can help meteorologists make data-driven decisions related to disaster preparedness, risk assessment, and resource allocation based on the expected intensity of different weather events.



2. Are there any trends in the intensity of storms impacting the Atlantic Basin?

As per the data This Map shows all categories and Hurricanes, Tropical depression, Tropical storms happened from 1975-2021 for the Atlantic Basin. The Pins shows where the storms started and how many happened in that selected year from the filter.

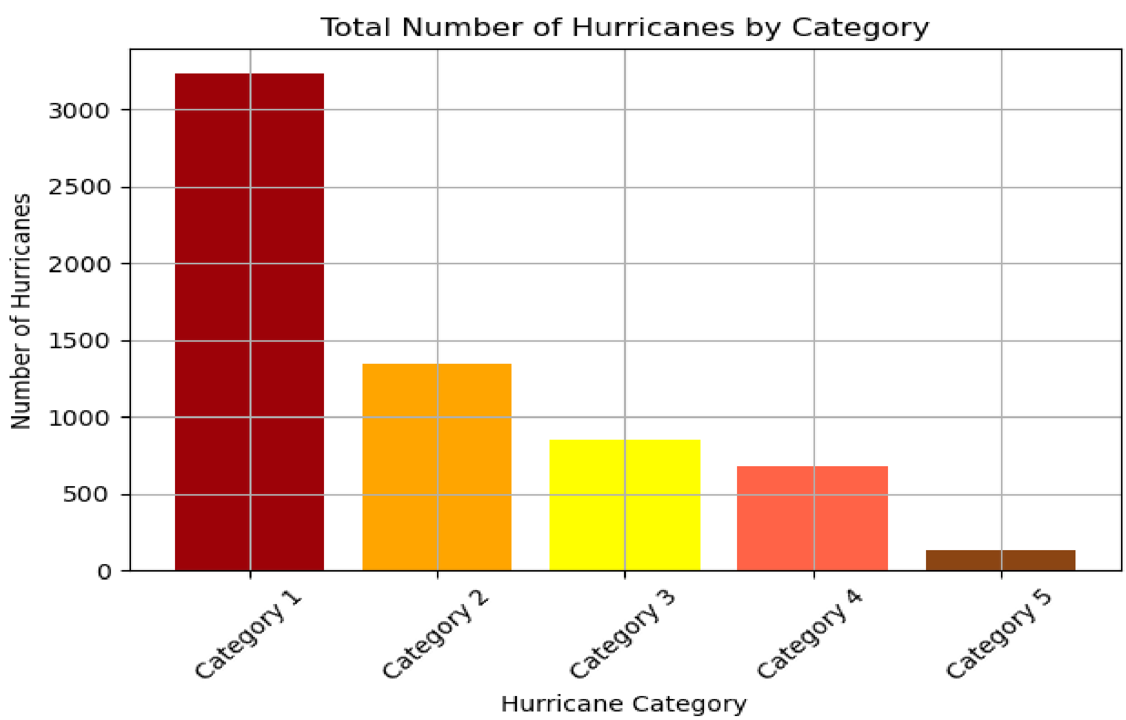
- By mapping the different storm categories (Hurricanes, Tropical Depressions, Tropical Storms) that occurred in the Atlantic Basin during this period, you can observe the distribution of storm intensities. This can help identify regions where more intense storms tend to form.
- Tracking the number of storms that occurred each year within the selected timeframe can show temporal trends in storm activity. You may observe fluctuations in the frequency of storms over the years, including periods with increased or decreased storm activity.
- The map pins showing where the storms started can highlight specific areas or regions that are prone to storm formation. Identifying these hotspots can be valuable for understanding the geographical patterns of storm genesis in the Atlantic Basin.
- Changes in climate factors such as sea surface temperatures, atmospheric conditions, and ocean currents can influence the intensity and frequency of storms in the Atlantic Basin. Analyzing the storm data over several decades can provide insights into how these climate variations impact storm activity.
- Hurricane seasons in the Atlantic Basin typically peak in late summer and early fall. Observing the distribution of storm categories throughout the hurricane season can reveal seasonal patterns in storm intensity.
  By visualizing the storm data on a map and tracking storm categories and locations from 1975 to 2021, you can uncover trends in storm intensity, identify regions prone to storm formation, and gain insights into the historical patterns of storm activity in the Atlantic Basin.
- The insights from the map can help meteorologists make data-driven decisions related to disaster preparedness, risk assessment, and resource allocation based on the expected intensity of different weather events.


Team Storm
2.0
Home
Dashboard
Map
About Us
Work Cited
2020
United Kingdom
Беларусь
Deutschland
Україна
France
România
Қазақстан
Italia
O'zbekiston
España
Ελλάς
Türkiye
Türkmenistan
United States
Algérie
México
India
Mali
Niger
Tchad
Nigeria
South Sudan
Leaflet | © OpenStreetMap contributors
Enter your email to subscribe...
Subscribe

3. Are there any trends in the appearance or movement of Hurricanes vs Categories?

As per the data No of Hurricanes vs Category shows Category 1 is the most common hurricane occurrence and as the hurricane category increases the chances of a hurricane of that category occurring decreases.

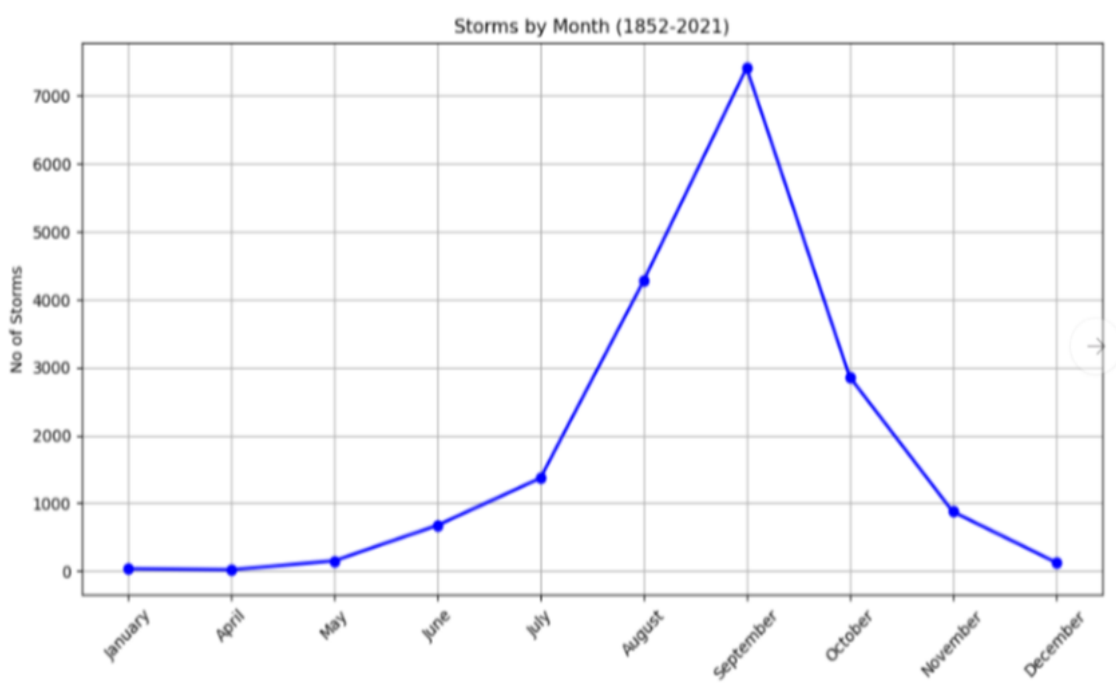
- One of the most notable trends is the increase in wind speed and intensity as the hurricane category rises. Category 1 hurricanes have lower wind speeds compared to Category 5 hurricanes, which are known for their extremely high wind speeds and destructive potential.
- Category 1 hurricanes are more common than Category 5 hurricanes. This trend of decreasing frequency as the category increases is consistent with the scale's design to reflect the rarity of the most intense hurricanes.
- The insights from the chart can help meteorologists make data-driven decisions related to disaster preparedness, risk assessment, and resource allocation based on the expected intensity of different weather categories.



4. Which seasons affect the Storms over the period (1852-2021)?

   As per the data the No of Storms vs Months shows exponential growth starting in the season of spring until it reaches its apex in September during summer and exponentially decreases going into the fall and winter seasons.

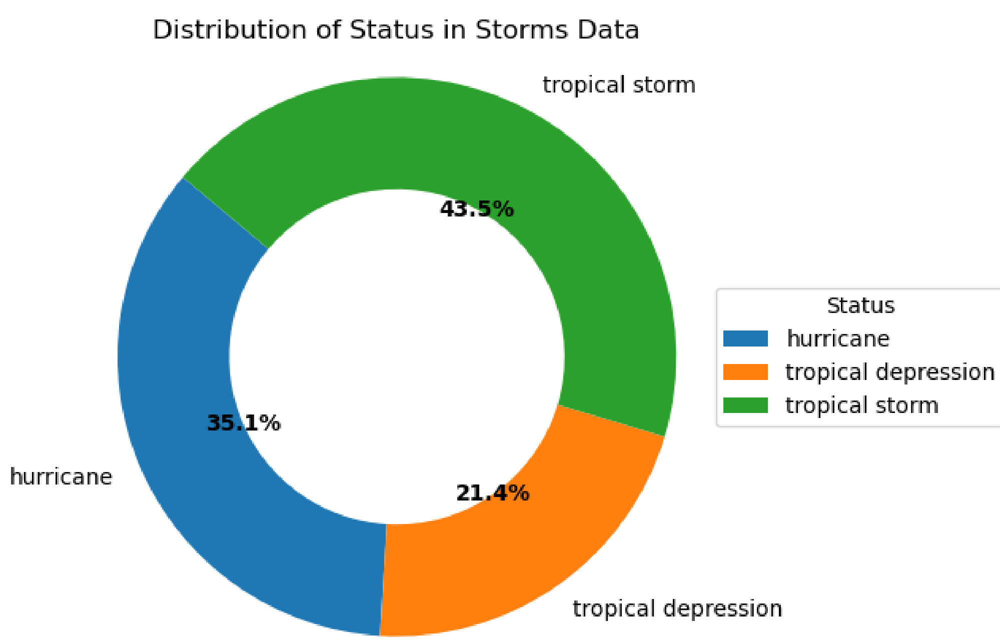
- Introduction: Analysis of storm frequency by month reveals seasonal patterns that influence storm occurrence over the period from 1852 to 2021.
- Summer Season: Storm frequency shows an exponential growth starting in the spring season, indicating an increase in storm activity.
- Fall Season: The peak of storm activity is observed in September during the fall season, with the highest number of storms occurring during this period.
- Fall and Winter Seasons: Storm frequency exponentially decreases as the seasons transition from Fall to Winter, indicating a decline in storm activity.
- The insights from the chart can help meteorologists make data-driven decisions related to disaster preparedness, risk assessment, and resource allocation based on the expected intensity of different weather seasons.



5. Donut chart with distribution of Hurricanes, Tropical Depression, Tropical Storms.

As per the data, the donut chart shows Hurricanes, Tropical Depression, and Tropical Storms. The chart shows the percentage of the whole over a period.

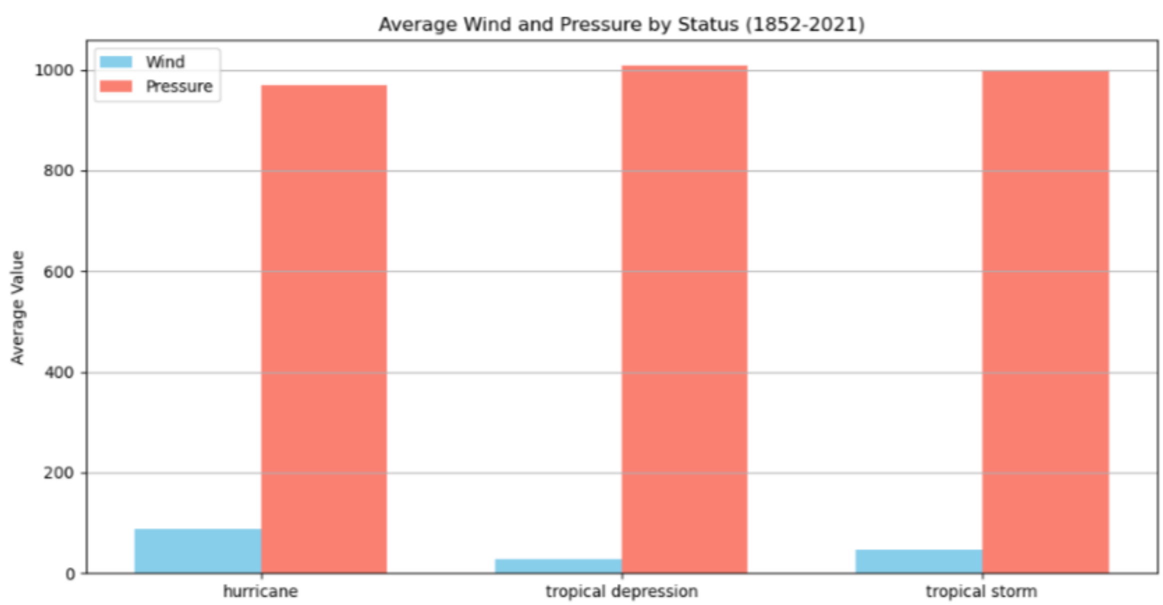
- The donut chart is divided into three segments, where each segment represents a category of Hurricanes, Tropical Depression, and Tropical Storms.
- In the Atlantic region, 43.5% of the most severe tropical storms are having a significant impact.
- In the Atlantic region, tropical depressions account for 21.4% of the weather events based on the donut chart data.



6. Wind and Pressure vs Status 1852-2021.

As per the data, the bar chart provides a visual representation of the average wind speed and pressure for various weather categories, shows valuable insights into the intensity, trends, and potential impact of hurricanes, tropical depressions, and tropical storms over the specified period.

- The chart shows a comparison of the average wind speeds and pressures associated with hurricanes, tropical depressions, and tropical storms. This comparison can provide insights into the relative intensity of these weather events.
- By observing the trends in wind speed and pressure over the years for each category, viewers can identify any patterns or changes in the intensity of these weather events over time.
- Understanding the average wind speed and pressure for each weather category can provide insights into the potential impact of these events. Higher wind speeds and pressures may indicate more severe weather conditions and potential risks.
- The insights from the chart can help meteorologists make data-driven decisions related to disaster preparedness, risk assessment, and resource allocation based on the expected intensity of different weather events.



Call to Action:

By analyzing all the charts and maps Generally, higher category hurricanes are more destructive and intense compared to lower category hurricanes. Category 5 hurricanes, for example, are known for their extreme wind speeds and potential for catastrophic damage. Category 1 hurricanes are more common than higher category hurricanes. This trend indicates that while stronger hurricanes are less frequent, they can have a significant impact when they do occur. Higher category hurricanes tend to be more closely monitored due to their potential for causing widespread damage. These hurricanes may also follow specific paths that can impact coastal regions differently based on their category and trajectory. With the effects of climate change, there is ongoing research into how hurricanes may be affected in terms of frequency, intensity, and behavior. This could potentially influence future trends in hurricane occurrence and movement patterns. The Atlantic hurricane activity is known to exhibit cyclical patterns, such as the Atlantic Multidecadal Oscillation (AMO). This natural climate cycle can influence the number and intensity of hurricanes over multiple decades. Observing these cycles can help identify periods of increased or decreased hurricane intensity. Different regions within the Atlantic Basin may experience varying trends in hurricane intensity. Some areas may be more prone to major hurricanes, while others may see fewer intense storms. Mapping the intensity of storms over the years can highlight these regional differences. The seasons, wind, pressure and categories and analyzing the map will help you understand the path of the storms.

By analyzing historical hurricane data and tracking current storms, researchers can continue to study these trends to better understand and prepare for the impact of hurricanes of different categories.

Limitations/Biases:

While the dataset provides valuable information on storms in the Atlantic Basin from 1852 to 2021, the lack of data from other regions may restrict the ability to draw comprehensive comparisons or analyze global storm patterns. Satellites were not widely available during that time, and some storms may not have been reported.

- When working with a dataset that has such limitations, it's essential to be transparent about these constraints in any analysis or conclusions. Additionally, considering the dataset's scope and focusing on insights specific to the Atlantic Basin can help mitigate the impact of these limitations on the overall analysis. If

needed, future analyses could incorporate data from other regions to provide a more comprehensive understanding of storm patterns globally.
- The data may be biased towards certain geographical regions that are more prone to storms, leading to underrepresentation of storms in less affected areas.
- The dataset may not include all storms or may have missing data for certain variables, which can affect the analysis and conclusions drawn from the data.
- The data may be biased towards more recent storms or seasons, potentially overlooking historical storm patterns and trends.
- Errors in data collection or recording could introduce inaccuracies in the dataset, impacting the reliability of the analysis.
- The dataset may only include storms of a certain magnitude or type, excluding others and leading to a skewed representation of storm characteristics.
- Storm data may be influenced by reporting biases, where certain types of storms are more likely to be reported or recorded than others, affecting the overall dataset's completeness.

Severe storms can have significant impacts on communities, causing damage to infrastructure, disrupting essential services, and posing risks to human life. To address these challenges, it's important for communities to have effective emergency preparedness plans in place. This includes early warning systems, evacuation procedures, shelters, and communication strategies to keep residents informed and safe during severe weather events.

Additionally, data visualization can play a crucial role in understanding and preparing for severe storms. By visualizing weather data, such as storm tracks, intensity, and potential impact areas, emergency responders and decision-makers can better plan and allocate resources to mitigate the effects of severe weather events.

Future work:

On this Storms data we have very limited columns/variables to work with to analyze how the storms are impacting surrounding areas other than the Atlantic Basin region. If we need to expand our analysis to include data from other regions, we want to consider accessing global storm data sources such as the World Meteorological Organization (WMO), the National Oceanic and Atmospheric Administration (NOAA), or other meteorological agencies around the world.

By incorporating data from various regions, you can gain a more comprehensive understanding of storm patterns, trends, and impacts on a global scale. This broader dataset can provide valuable insights into storm behavior, frequency, intensity, and potential future trends.

When working with storm data from different regions, it's important to ensure that the data sources are reliable, consistent, and properly documented. Additionally, consider the specific variables and parameters you want to analyze, such as storm tracks, wind speeds, precipitation levels, and any other relevant factors.

Additional columns/variables to analyze, duration of storm, casualties and damage caused by tropical storms and hurricanes, focus on other natural disasters such as wildfires and earthquakes.

Work Responsibilities:

Responsibilities were shared between the team members Thripura did the Project proposal, Eugenio Elizondo created the GitHub project repository, Lindsey and Thripura did the daily standup communication. We did the data cleaning as a Team, as we chose to work with Storms data with all numerical values and data has no null values, mismatching values, no duplicates so data cleaning was easy with this dataset and created a 1st jupyter notebook.
Mason worked on a 2nd jupyter notebook to create a database. SQLite, and sqlhelper.py with SQL alchemy code for the visuals. Hassan worked to create app.py and HTML pages and Json files. Visualization part Lindsey worked on a Stacked Bar chart, Thripura created line chart, Eugenio is diligently tackling a complex map visualization and a bar chart. Lindsey is working on PPT and Thripura working on Writeup, Team will review and adjust as necessary. We all worked very closely on the project coordination, each taking the lead when needed. Assistance from TAs as needed occurred during the project as well as suggestions from the professor.

Sources:

Our Website Link: https://56f7-2603-8081-4e06-7f4-e9c7-561e-395-47b2.ngrok-free.app/

Data Set: https://www.kaggle.com/datasets/christinezinkand/storms

Slide Deck Template: https://slidesgo.com/theme/biology-subject-for-high-school-natural-disasters-tropical-cyclone#position-14&related-1&rs=detail-related

https://airandspace.si.edu/collection-objects/tiros-meteorological-satellite/nasm_A19650289000#:~:text=Brief%20Description,features%20for%20the%20first%20time

https://scijinks.gov/hurricane/

Took lots of help with: Xpert Learning

https://journals.ametsoc.org/view/journals/mwre/141/2/mwr-d-12-00152.1.xml

Color Palette: https://www.color-hex.com/color-palette/114111