A close-up of a logo

Description automatically generated

Submitted by: Group 40

Name: Pranav Vakadikar

UB No: 22051306

Name: Danyal Yasin

UB No: 22052309

Name: Vipanjot Kaur

UB No: 22068901

**COS7046-B: Big Data Visualisation**

Table of Contents

[Introduction: 3](#_Toc155810919)

[Objectives: 3](#_Toc155810920)

[Background: 3](#_Toc155810921)

[Visualization technologies and platforms: 3](#_Toc155810922)

[Relevant approaches and systems: 3](#_Toc155810923)

[Main Part: 4](#_Toc155810924)

[Explanation of the space weather visualization problem: 4](#_Toc155810925)

[Demonstration of visualization techniques: 4](#_Toc155810926)

[Analysis and discussion: 8](#_Toc155810927)

[Conclusion: 9](#_Toc155810928)

[References: 9](#_Toc155810929)

**Table of figures :**

[Figure 1: Distribution of Solar Flares 4](#_Toc155810961)

[Figure 2 : Top 20 Solar Flare Type 5](#_Toc155810962)

[Figure 3: Distribution of Flare Type 5](#_Toc155810963)

[Figure 4:Distribution of number of sunspots 5](#_Toc155810964)

[Figure 5:Figure 5: Distribution of penumbra 6](#_Toc155810965)

[Figure 6:Types of compactness 6](#_Toc155810966)

[Figure 7: Distribution of magnetic class 6](#_Toc155810967)

[Figure 8: Top 10 stations 7](#_Toc155810968)

[Figure 9: Distribution of top 10 type of solar flare type 7](#_Toc155810969)

[Figure 10: Distribution of location 7](#_Toc155810970)

[Figure 11: Co relation matrix for direction 8](#_Toc155810971)

[Figure 12: Time series data for year 2016 8](#_Toc155810972)

[Figure 13: Time series data from 2009 to 2018 8](#_Toc155810973)

# Introduction:

Solar flares and sunspots pose significant challenges in space weather forecasting. The occurrence of large solar flares can have a substantial impact on space weather, affecting Earth's environment and climate (Kusano et al., 2020). Additionally, the solar activity during solar cycle 24 has been linked to climate variations, highlighting the broader implications of solar activity on Earth's climate (Ansor et al., 2023). These references collectively underscore the complex interplay between solar flares, sunspots, and their impact on space weather, emphasizing the need for advanced forecasting models and mitigation strategies to address these challenges.

# Objectives:

The understanding of solar flares and sunspots is crucial due to their significant impact on space weather and Earth's environment. Research by emphasizes the role of shearing motion and sunspot rotation in the build-up of free energy, leading to the formation of flux ropes in the corona, which produce solar flares and coronal mass ejections (Yan et al., 2018). Additionally, the study by underscores the importance of factoring in regional differences of sunspots in solar flare forecasting, validating the effectiveness of their approach (Fu et al., 2023). The research aims to uncover the correlation between sunspots and solar flares, enhancing our understanding of sunspot influence on solar phenomena. Visualizations will be developed to depict this correlation clearly, facilitating an intuitive grasp of associated patterns. Additionally, the project seeks to optimize prediction accuracy by selectively identifying crucial features in the dataset, streamlining it for a more efficient and precise model that anticipates solar flare events based on sunspot activity.

# Background:

## Visualization technologies and platforms:

The background of solar flare and sunspot visualization involves harnessing advanced technologies and platforms, with Python playing a pivotal role. The development of visualization technologies and platforms has significantly transformed the analysis of space weather data. Integration of visualization tools as "models" that can be coupled with other integrated models enables comprehensive analysis (Poedts et al., 2020). Furthermore, Python, with its rich ecosystem of libraries such as Matplotlib and SunPy, along with the utilization of new computational techniques like machine learning and numerical methods, facilitates unravelling information from expanding datasets through advanced visualizations (Fogg et al., 2023). These references underscore the critical role of visualization technologies, including Python, in enhancing our understanding of space weather and its impacts on human technology and infrastructure.

## Relevant approaches and systems:

Relevant approaches and systems for solar flares and sunspots encompass a diverse range of studies. One innovative system involves developing a machine learning model that associates sunspots with solar flares, utilizing publicly available solar catalogues. This approach, inspired by the work of Qahwaji and Colak (2007), aims to establish a correlation between the occurrence of sunspots and the likelihood of solar flares. By leveraging visualization techniques, the system analysis patterns within solar data, providing an insightful framework for anticipating solar flare events based on the presence of sunspots. This methodology not only contributes to understanding the interdependence of these solar phenomena but also offers a practical tool for space weather forecasting, showcasing the intersection of data science and solar physics in addressing challenges related to solar flares and sunspots.

# Main Part:

## Explanation of the space weather visualization problem:

The visualization of space weather poses a complex challenge, as it necessitates the comprehensive monitoring, analytical control, and visualization of various parameters such as geomagnetic field variations and solar activity, as outlined by Воробьев & Vorobeva (2016). The goal is to offer tangible visual aids to space weather experts, enhancing their ability to make accurate forecasts, as highlighted by Lapenta et al. (2013). The intricate nature of space weather data requires sophisticated visualization techniques to discern patterns, trends, and anomalies in real-time. This task involves not only representing the spatial distribution of solar phenomena but also conveying the temporal evolution of these events. Achieving this balance is crucial for producing effective visualizations that support space weather experts in deciphering and predicting the dynamic conditions of the space environment.

## Demonstration of visualization techniques:

**Q1**: What is the most prevalent flare type observed in the dataset, and how does its frequency compare to other flare classifications?

A screen shot of a computer

Description automatically generated

Figure 1: Distribution of Solar Flares

A screenshot of a computer

Description automatically generated

Figure 2 : Top 20 Solar Flare Type

A screenshot of a computer

Description automatically generated

Figure 3: Distribution of Flare Type

The depicted bar graph illustrates the flare distribution based on their types, showcasing the top 20 flare types. It is evident that C11 flares are the most prevalent, with C10, C12, and C13 following closely. The third plot indicates that C-type flares are the most frequently occurring, with B-type flares ranking next in prevalence. M-type flares are notably less common, while X-type and A-type flares are observed to be exceedingly rare.

**Q2**: What is the most common type of sunspot observed in the dataset, and how does its associated penumbra and compactness types compare to other classifications?

A screenshot of a computer

Description automatically generated

Figure 4:Distribution of number of sunspots

A screenshot of a computer

Description automatically generated

Figure 5:Figure 5: Distribution of penumbra

A screenshot of a computer

Description automatically generated

Figure 6:Types of compactness

Examining the graph, it is evident that B-type sunspots are the most prevalent, followed by A-type. Additionally, some values are missing, indicating spots that cannot be classified. In the second plot, depicting penumbra types, S, A, and X types are nearly equally common, while K, R, and H types occur less frequently. The third bar graph illustrates compactness types, with O being the most common, followed by X and I. Types C are among the least occurring, along with null values and 0.

**Q3**: What are the ten stations that have gathered the most data in the dataset, and how does the volume of data collected by these stations compare to others in the study?

A screenshot of a computer

Description automatically generated

Figure 7: Distribution of magnetic class

A screenshot of a computer

Description automatically generated

Figure 8: Top 10 stations

A screenshot of a computer

Description automatically generated

Figure 9: Distribution of top 10 type of solar flare type

The presented graphs highlight the top 10 stations acquiring the most data, with LEAR being the station that has acquired the highest amount of data. In the graph below, depicting the 10 most prevalent magnetic classes, B-type is the most frequently occurring class.

**Q4** : How are sunspots distributed in terms of their location along the North-South and East-West axes, and corelation in the dataset?

A screenshot of a computer

Description automatically generated

Figure 10: Distribution of location

A screenshot of a computer

Description automatically generated

Figure 11: Co relation matrix for direction

**Q5** : Time series data that shows sun spots from 2009 to 2019 and for the year of 2016 A screen shot of a computer

Description automatically generated

Figure 12: Time series data for year 2016

A screenshot of a computer

Description automatically generated

Figure 13: Time series data from 2009 to 2018

# Analysis and discussion:

The comprehensive analysis of the presented data reveals crucial insights into the most prevalent types across various solar phenomena. Examining sunspots, the predominant occurrence of B-type sunspots stands out prominently, with A-type sunspots closely following suit. The prevalence of missing values suggests challenges in classifying certain spots, warranting further investigation. In the context of penumbra types, the nearly equal frequency of S, A, and X types indicates their significance in the dataset. K, R, and H types, while occurring less frequently, contribute to the overall diversity.

Delving into compactness types, the prominence of O-type is noteworthy, followed by X and I type. Conversely, types C are among the less occurring, along with null values and 0 in the dataset, illustrating the diversity of compactness characteristics.

Shifting the focus to data acquisition, the top 10 stations with LEAR at the forefront have been instrumental in gathering the most data, underscoring its significance in data acquisition efforts. In terms of magnetic classes, the graph underscores the prevalence of B-type as the most frequently occurring class, providing a crucial answer to the overarching question of the most relevant types across sunspots, penumbra, compactness, and magnetic classes in the dataset. This synthesis of findings not only enhances our understanding of solar phenomena but also contributes to refining predictive models for more accurate space weather forecasting.

# Conclusion:

In conclusion, the analysis of penumbra distribution and compactness types in solar datasets provides valuable insights into the prevalence of specific solar phenomena. The observed patterns, such as the dominance of S, A, and X types in penumbra distribution and the frequent occurrence of O, X, I, and C types in compactness, serve as essential benchmarks for understanding solar activity. In the distribution of solar flares, C11 flares are the most common, followed closely by C10, C12, and C13. C-type flares dominate, with B-type flares ranking second, while M-type flares are less frequent. X-type and A-type flares are notably rare. On the other hand, in the distribution of sunspot types, B-type sunspots are the most prevalent, followed by A-type. There are also instances of missing values, indicating unclassified spots. In terms of magnetic classes, B-type is the most occurring, highlighting its prevalence among the top 10 magnetic classes These findings are particularly relevant when considering the most prevalent sunspots and flares, as they contribute to our broader understanding of space weather dynamics. The identification of rare classifications, such as O-type penumbra and the minimal representation of O-type compactness, prompts further exploration and consideration within the broader solar context. These results not only inform ongoing space weather research but also lay the groundwork for refining predictive models. Moving forward, delving into the intricate relationships between these variables and expanding the analysis across varied temporal and spatial scales will be crucial for advancing our comprehension of the dynamic processes occurring on the Sun and improving space weather forecasting capabilities.

# References:

1. Ansor, N., Johari, N., Hamidi, Z., & Shariff, N. (2023). Solar activity-climate relations during solar cycle 24. Iop Conference Series Earth and Environmental Science, 1151(1), 012022. https://doi.org/10.1088/1755-1315/1151/1/012022
2. Fogg, A., Jackman, C., Malone-Leigh, J., Gallagher, P., Smith, A., Lester, M., … & Waters, J. (2023). Extreme value analysis of ground magnetometer observations at valentia observatory, ireland. Space Weather, 21(7). https://doi.org/10.1029/2023sw003565
3. Fu, J., Wan, J., Zhao, Z., Han, K., & E, P. (2023). Federated transfer learning for soalr flare forecasting.. https://doi.org/10.21203/rs.3.rs-3134736/v1
4. Kusano, K., Iju, T., Bamba, Y., & Inoue, S. (2020). A physics-based method that can predict imminent large solar flares. Science, 369(6503), 587-591. https://doi.org/10.1126/science.aaz2511
5. Lapenta, G., Poedts, S., & Vucinic, D. (2013). Space weather prediction and exascale computing. Computing in Science & Engineering, 15(5), 68-76. https://doi.org/10.1109/mcse.2012.86
6. Poedts, S., Kochanov, A., Lani, A., Scolini, C., Verbeke, C., Hosteaux, S., … & Luntama, J. (2020). The virtual space weather modelling centre. Journal of Space Weather and Space Climate, 10, 14. https://doi.org/10.1051/swsc/2020012
7. Qahwaji, R. and Colak, T. (2007). Automatic short-term solar flare prediction using machine learning and sunspot associations. Solar Physics, 241(1), 195-211. https://doi.org/10.1007/s11207-006-0272-5
8. Yan, X., Wang, J., Pan, G., Kong, D., Xue, Z., Yang, L., … & Feng, X. (2018). Successive x-class flares and coronal mass ejections driven by shearing motion and sunspot rotation in active region noaa 12673. The Astrophysical Journal, 856(1), 79. https://doi.org/10.3847/1538-4357/aab153
9. Воробьев, А. and Vorobeva, G. (2016). Analytical information system for control and spectral analysis of geomagnetic field and space weather parameters. Russian Journal of Earth Sciences, 16(4), 1-9. https://doi.org/10.2205/2016es000571