Seasonal Variation in Aging-Associated Health Measures: Alzheimer's and Mental Health Patterns

018202600 Basanth Periyapatna Roopa Kumar , 018202106 Kushal Atulbhai Adhyaru
018190003 Nandhakumar Apparsamy , 018174104 Nischitha Nagendran , 017777045 Prem Shah

***Abstract*—Aging populations around the world face distinct health difficulties, with Alzheimer's disease and mental health issues emerging as key areas of concern. This research investigates seasonal variations in aging-related health measurements to uncover dynamic shifts in health patterns throughout time. Using data from the Alzheimer's Disease and Healthy Ageing datasets, we investigate demographic differences, spatial variances, and their combined impact on public health outcomes. This study uses interactive Power BI dashboards and delivers actionable data for healthcare planners and policymakers by highlighting geographical and demographic differences in health measures. The findings reveal tendencies in high-risk populations and help drive focused actions to enhance aging-related health outcomes. The visualizations, which include demographic breakdowns, geographic heatmaps, and yearly trends, provide a fascinating narrative of the changing environment of aging and health in the United States. This study offers decision-makers a comprehensive overview of aging health measures, promoting data-driven strategies to reduce inequities and improve community well-being. We hope to transform public health planning for aging populations by combining data and impactful visualization.**

***Index Terms*—Alzheimer's Disease, Healthy Aging, Data Visualization, Power BI Dashboards, Public Health Insights, Interactive Visualizations, Healthcare Decision-Making, Predictive Analytics, Data Cleaning**

I. Introduction

THIS project is intended to serve as a key component in the health sector. As the world's population ages, the convergence of public health and aging creates both possibilities and difficulties for healthcare systems and governments. Alzheimer's disease and mental health issues have arisen as major focal points, necessitating a better knowledge of the complex health metrics that influence these disorders. However, the amount of fragmented health data and the lack of intuitive systems for comprehensive analysis impede practical insights. This study addresses these complications by developing a comprehensive, visualization-based strategy for identifying trends in aging-related health metrics.

Using the Alzheimer's Disease and Healthy Ageing dataset, this study applies an advanced data visualization approach to identify spatial and demographic trends that influence health outcomes. The project's goal is to give stakeholders a unique toolbox for analyzing high-risk groups and identifying healthcare disparities by integrating interactive Power BI dashboards. By moving beyond static data representations, this study shows dynamic spatial inequities, demographic stratifications, and temporal health patterns.

This paper was produced by Group 3 Team for Data Visualization Course

The remaining sections of this report are arranged as follows. Section II examines relevant literature, emphasizing past research on aging health measures and data visualization tools. Section III details the approach, which includes problem conceptualization, data discovery, preparation, visualization modeling, evaluation, and deployment. Section IV summarises the findings and insights gleaned from the visualizations. Section V outlines the technological problems faced throughout the project. Section VI discusses lessons learned, and Section VII closes the study with recommendations for future work. Finally, Section VIII recognizes the efforts that aided this research.

II. Literature Review

A. Study I: Evaluating the Alzheimer's Disease Data Landscape

Birkenbihl et al. conducted an extensive evaluation of the Alzheimer's disease (AD) data landscape, analyzing patient-level data from nine major clinical cohort studies. This research revealed critical demographic biases, including an over-representation of White/Caucasian individuals, and emphasized the challenges in data interoperability due to inconsistent variable naming and diverse data models. The study also emphasized the scarcity of longitudinal biomarker data, such as cerebrospinal fluid samples, which are critical for determining illness progression. The researchers made significant discoveries more accessible by creating ADataViewer, an interactive web tool for analyzing these datasets. The systematic approach to comparing and characterizing cohort datasets underscores the importance of meticulous data evaluation for generating robust and reproducible insights. This work informed our emphasis on identifying disparities and biases within demographic and geographic data in our project, particularly in visualizing patterns related to Alzheimer's disease and mental health. The emphasis on systematic data assessment and the challenges of interpreting diverse datasets resonated with our methodology for designing dynamic visualizations tailored to healthcare decision-making.

B. Study II: Visualization of Alzheimer's Disease Progression in Low-Dimensional Manifolds

Seo et al. (2019) suggested a unique method for tracking Alzheimer's disease (AD) progression that employs nonlinear dimensionality reduction techniques, specifically Locally Linear Embedding. They used longitudinal MRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) to map

AD progression across time on a low-dimensional manifold. The study emphasized the importance of LLE in retaining local structures in high-dimensional data and improved diagnostic visualization using a two-dimensional LLE map. When combined with Support Vector Machines (SVM), the approach demonstrated great accuracy in categorizing cognitively unimpaired (CU) and AD patients while also allowing for longitudinal tracking of individual disease development. The study emphasizes the potential of nonlinear approaches to improve AD diagnosis and monitoring.

C. Study III: Machine Learning-Driven Analysis of Alzheimer's Research Trends and Hotspots

A study by Guan et al. utilized machine learning techniques to analyze research trends and hotspots in Alzheimer's disease (AD) using PubMed abstracts. This study identified long-term research hotspots, including diseases, proteins, and symptoms, while also uncovering evolving trends, such as an increased focus on hormones and amyloid seed contamination in 2016. The use of modern data processing and clustering approaches in their work emphasizes the significance of carefully analyzing complicated datasets in order to disclose useful insights. While their focus was on AD research trends, their approach to identifying patterns and trends influenced our methods for visualizing demographic and geographic health disparities connected to Alzheimer's disease and mental health patterns. The emphasis on discovering patterns in complicated data is consistent with our goal of providing insights to healthcare planners via dynamic visualization frameworks.

D. Study IV: Mapping the Alzheimer's Disease Research Landscape with Network and Content Analysis

Song et al. adopted a novel approach to explore the landscape of Alzheimer's disease (AD) research, integrating concept graph-based network analysis and topic modeling to analyze metadata from 96,081 articles in PubMed. By constructing graphs of biological entities and their semantic relationships, they identified key nodes, such as amyloid-beta protein and tau protein, and analyzed their significance using centrality measures. In addition, 16 important research themes were identified, with a noticeable increase in the emphasis on transgenic mouse models and improved imaging techniques over time. Unlike standard bibliometric analyses, this study gave deeper insights by connecting bio-entities to developing trends, providing a roadmap for future AD research orientations. The study's thorough integration of macro and micro-level studies motivated us to focus on contextualizing trends in varied datasets to generate practical insights for healthcare and policymaking.

E. Study V: Evolution of Alzheimer's Disease Research through Text Mining

Martinelli (2022) used text mining techniques such as frequency analysis and LDA topic modeling to examine roughly 17,000 abstracts from Alzheimer's research, revealing insights into disease processes, treatment options, and interdisciplinary collaboration trends. This work highlighted the utility of computational techniques for synthesizing vast amounts of research data, discovering new trends, and facilitating cross-disciplinary discovery. The innovative use of text mining to extract actionable knowledge inspired us to concentrate on leveraging dynamic visualization and health-tech breakthroughs to identify trends in Alzheimer's and mental health data.

III. Implementation Flow

The implementation of the project was categorized into four critical phases, each consciously designed to assure progress and results in a structured manner. In this regard, the flow of implementation is presented below in detail.

A. Charting the Course (September 2024)

The first phase was dedicated to understanding the scope and context of the project.

Objective: To clearly define the direction and lay a concrete foundation for studying aging-related health metrics through interactive visualizations.

Steps Taken:

- Extensively researched Alzheimer's Disease and mental health issues prevalent in aging populations.
- The review of the literature and datasets was performed with the objective of outlining the key challenges and gaps in research.
- Defined the problem statement - the need for actionable insights to help policymakers in planning healthcare.

Outcome: It laid the foundation so that the project was focused in approach, with the objectives aligning with public health priorities.

B. Laying the Groundwork (October 2024)

This stage concerned the acquisition, cleaning, and preparation of the dataset for analysis.

Objective: The preparation of the Alzheimer's Disease and Healthy Ageing dataset for exploratory analysis and visualization.

Steps Taken:

- Acquired the dataset and performed rigorous data cleaning, including:
- Removal of unimportant columns and imputation of missing values.
- Standardizing data formats for consistency among demographic, spatial, and seasonal fields.
- Performed exploratory data analysis to highlight preliminary patterns, such as trend analyses of the mental health and aging metrics.
- Developed foundational visualizations using Power BI to highlight demographic disparities and seasonal patterns.

Outcome: The dataset was transformed into a clean and structured format, allowing for the creation of meaningful early-stage insights.

C. Midway Checkpoint (November 2024)

The midway phase emphasized refining visualizations and enhancing interactivity.

Objective: To develop advanced visualizations and collect stakeholder feedback to improve dashboard utility.

Steps Taken:

- Iteratively developed visualizations, integrating advanced Power BI features, such as:
  - Slicers and filters for user-driven exploration of data.
  - Drill-down capabilities for detailed demographic and spatial analysis.

• Presented prototype dashboards to stakeholders, including public health policymakers, for feedback. • Improved visualizations incorporating review comments to ensure clarity, accessibility, and actionable insight. • Focused analysis was done on specific trends; for example, the impact of seasonality on Alzheimer's and mental health outcomes.

The dashboards became dynamic and highly interactive, providing actionable insights to stakeholders in decision-making.

D. Destination Impact (December 2024)

The final phase resulted in the delivery of a polished and comprehensive Power BI dashboard.

Objective: Showcase a functional and impactful visualization tool to support healthcare planning.

Steps Taken:

- Delivered a comprehensive Power BI dashboard featuring:
  - Interactivations visualizing seasonal and demographic trends of health metrics that change with age.
  - Actionable insights on reducing health inequity, improving health outcomes for an aging population.
- Developed a detailed presentation to summarize findings and display the key features of the dashboard.
- Engaged stakeholders with hands-on sessions to drive home the utility of this dashboard in real-world decision scenarios.
- Prepared documentation and a user guide for the seamless adoption of the tool.
- Outcome: It was a successful project because it met its goals, in that it empowered stakeholders by providing data-driven strategies with which to improve public health outcomes for aging populations.

Key Takeaways : It ensured that every milestone in a staged approach fit together for an impactful outcome. Balancing data preparation, iterative design, and stakeholder collaboration throughout the project delivered a robust solution to real-world healthcare challenges.

IV. Methodology

A. Using Python

1) Data Collection: Data for this project was gathered from various reliable sources to ensure that it was both accurate and comprehensive.

Primary Sources: Data repositories like the data.gov, including information on Alzheimer's disease prevalence, mental health indicators, and demographic data - age, gender, race/ethnicity.
Data scope: Multiple years and geographic regions covered variation in demographics and health conditions. This therefore allowed a stratified analysis to be performed by seasonality, age group, gender, race, and location.
The data collection process emphasized obtaining datasets that were consistent in structure and format to facilitate preprocessing and analysis.

2) Data Preprocessing: Preprocessing is a critical step to prepare raw data for analysis. The following detailed steps were undertaken:

2.2.1 Data Cleaning:

Null Value Handling: Missing values in the dataset were identified and treated using appropriate imputation strategies. For minor gaps, NaNs in columns such as `StratificationCategory2`, `Stratification2`, and `Geolocation` were filled with their respective mode (most frequent value). Columns with excessive missing data, including `Data_Value`, `Data_Value_Alt`, `Data_Value_Footnote_Symbol`, `Data_Value_Footnote`, `Low_Confidence_Limit`, and `High_Confidence_Limit`, were removed to maintain data integrity.

Outlier Detection: To ensure robust analysis, Z-score analysis and other statistical techniques were applied to detect and manage extreme values that could skew the results. This step was critical in preserving the reliability and accuracy of subsequent visualizations and insights.

Irrelevant Fields: Redundant or non-essential columns were removed to streamline the dataset and focus the analysis on relevant variables. For example, identifiers that did not contribute to the research questions were excluded, improving computational efficiency and clarity.

These preprocessing steps ensured the dataset was clean, consistent, and ready for meaningful analysis, setting the foundation for uncovering actionable insights.

Missing Data Overview

The following table summarizes missing data in selected columns:

TABLE I
MISSING DATA OVERVIEW

Column	Non-Null Count	Percentage Missing
Data_Value	192,808	32.14%
Low_Confidence_Limit	192,597	32.21%
High_Confidence_Limit	192,597	32.21%
StratificationCategory2	247,269	12.97%
Geolocation	253,653	10.73%

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 284142 entries, 0 to 284141
Data columns (total 25 columns):
 #   Column                      Non-Null Count   Dtype
---  ------                      --------------   -----
 0   RowId                       284142 non-null  object
 1   YearStart                   284142 non-null  int64
 2   YearEnd                     284142 non-null  int64
 3   LocationAbbr                284142 non-null  object
 4   LocationDesc                284142 non-null  object
 5   Datasource                  284142 non-null  object
 6   Class                       284142 non-null  object
 7   Topic                       284142 non-null  object
 8   Question                    284142 non-null  object
 9   Data_Value_Unit             284142 non-null  object
 10  DataValueTypeID             284142 non-null  object
 11  Data_Value_Type             284142 non-null  object
 12  StratificationCategory1     284142 non-null  object
 13  Stratification1             284142 non-null  object
 14  StratificationCategory2     284142 non-null  object
 15  Stratification2             284142 non-null  object
 16  Geolocation                 284142 non-null  object
 17  ClassID                     284142 non-null  object
 18  TopicID                     284142 non-null  object
 19  QuestionID                  284142 non-null  object
 20  LocationID                  284142 non-null  int64
 21  StratificationCategoryID1   284142 non-null  object
 22  StratificationID1           284142 non-null  object
 23  StratificationCategoryID2   284142 non-null  object
 24  StratificationID2           284142 non-null  object
dtypes: int64(3), object(22)
memory usage: 54.2+ MB
```

Fig. 1. The Data columns of the Dataset after cleaning the data

2.2.2 Data Duplication:

Duplicate records were identified by matching records across key fields of demographics, year, and health indicators. Duplicates were systematically removed to ensure that each row represented unique data points and thus preserved the integrity of the dataset.

2.2.3 Standardization:

Date and Time Formats: Temporal data was standardized to a consistent format, such as YYYY-MM-DD, to enable chronological analyses.

Categorical Encoding: Variables such as gender and race were encoded uniformly to eliminate discrepancies; for example, mapping ”M” and ”Male” to the same category.

Disaggregation: Data was disaggregated into subsets by year, location, and demographic factors such as age and gender. This allowed focused analyses and ensured that results aligned with the project’s goals.

3) Exploratory Data Analysis: EDA was a structured process of learning the nature and characteristics of the dataset in detail to draw some preliminary insights. The following types of analyses were done

Univariate Analysis:
Focused on individual variables like age, gender distribution, and frequency of Alzheimer’s diagnoses.
Visualizations such as histograms and pie charts depicted the distribution of these variables, highlighting key patterns. For example, higher prevalence in older age groups was portrayed.

Bivariate Analysis: Analyzed the relationships between two variables, including the correlation of mental health indicators with geographic regions.
Scatter plots were used to show trends, while boxplots provided insights into variations across groups, such as Alzheimer’s rates by gender.

Multivariate Analysis:
Analyzed interactions among multiple variables to uncover complex patterns. Techniques like clustering identified similar demographic groups, and PCA was used to reduce dimensionality while retaining essential trends.
Data Distribution and Insights:
Statistical summaries and visualizations, including heatmaps and violin plots, were employed to explore data distributions. Insights included the identification of peak seasons for mental health concerns and regional hotspots for Alzheimer’s prevalence.

B. PowerBI

1) Data Integration: The datasets were imported into Power BI from CSV files and online data repositories.

The data model was prepared in such a way that all tables were related to each other via some common fields, such as location and year, or demographic features.

2) Data Cleaning: Further cleaning was performed using Power Query Editor: Columns were renamed for better understanding. Filters were applied to exclude incomplete or irrelevant records. Duplicates were removed, and missing values were treated.

Transformation and Preparation:

Data transformation was done in the form of calculated fields, which supported specific analyses such as those required for seasonal averages and year-over-year changes.

Custom measures and calculated columns were developed with DAX to increase analytical depth.

3) Dashboard Design: Line Graphs: These visualized trends over time, such as the seasonal variation in Alzheimer's diagnoses and mental health issues.
Heatmaps: These showed geographic patterns and regional disparities in health metrics.
Bar Charts: Comparisons across demographics, including Alzheimer's prevalence by gender and race.
Drill-through Pages: Added for in-depth analysis, enabling users to click through and explore specific regions or years in detail.
Interactive Filters and Slicers: Users can refine the visualizations based on year, location, or demographic group.

4) Testing and Validation: The dashboard was thoroughly tested to ensure the following:

Accurate representation of data. Usability for end-users with varying technical skills.
Compatibility across devices and screen resolutions.

The final dashboard provided an interactive platform for exploring seasonal and demographic variations in health trends. The insights from the Power BI visualizations were developed for stakeholders, including policymakers, healthcare providers, and researchers, offering actionable knowledge to inform interventions and strategies.

C. Agile Methodology

We follow an Agile methodology to ensure flexibility, collaboration, and incremental progress. Agile focuses on iterative development, meaning that the project is divided into smaller, manageable pieces that are delivered in cycles called sprints. This enables continuous improvement and adaptation to changing requirements or insights gained during the project lifecycle. By involving stakeholders at every stage, Agile ensures that the outcomes align with user needs and the project objectives.

The Agile implementation in this project was done around four important stages:

Planning, Development, Review, and Refinement.

During the planning phase, the team had to clearly define the scope of the project, objectives, and tasks to be carried out and then prioritize them in a product backlog. Each sprint was focused on delivering specific components, such as data collection, preprocessing, EDA, or dashboard creation. Progress was tracked through daily stand-ups and bi-weekly sprint reviews to identify challenges and integrate feedback.

The core of the Agile process was collaboration. The project team, comprising data analysts, visualization experts, and stakeholders, works together to develop iterative deliverables. Regular feedback loops ensure the direction of the project remains relevant, stakeholders provide input on key deliverables, such as data insights or the usability of dashboards. During the final sprint, all components were integrated into a cohesive and polished solution to deliver an interactive Power BI dashboard to meet the project's objectives. This iterative collaborative framework was bound to maximize efficiency, foster innovation, and assure quality in delivered results.

V. Results

Fig. 2. Dashboard Home for navigating to different pages.

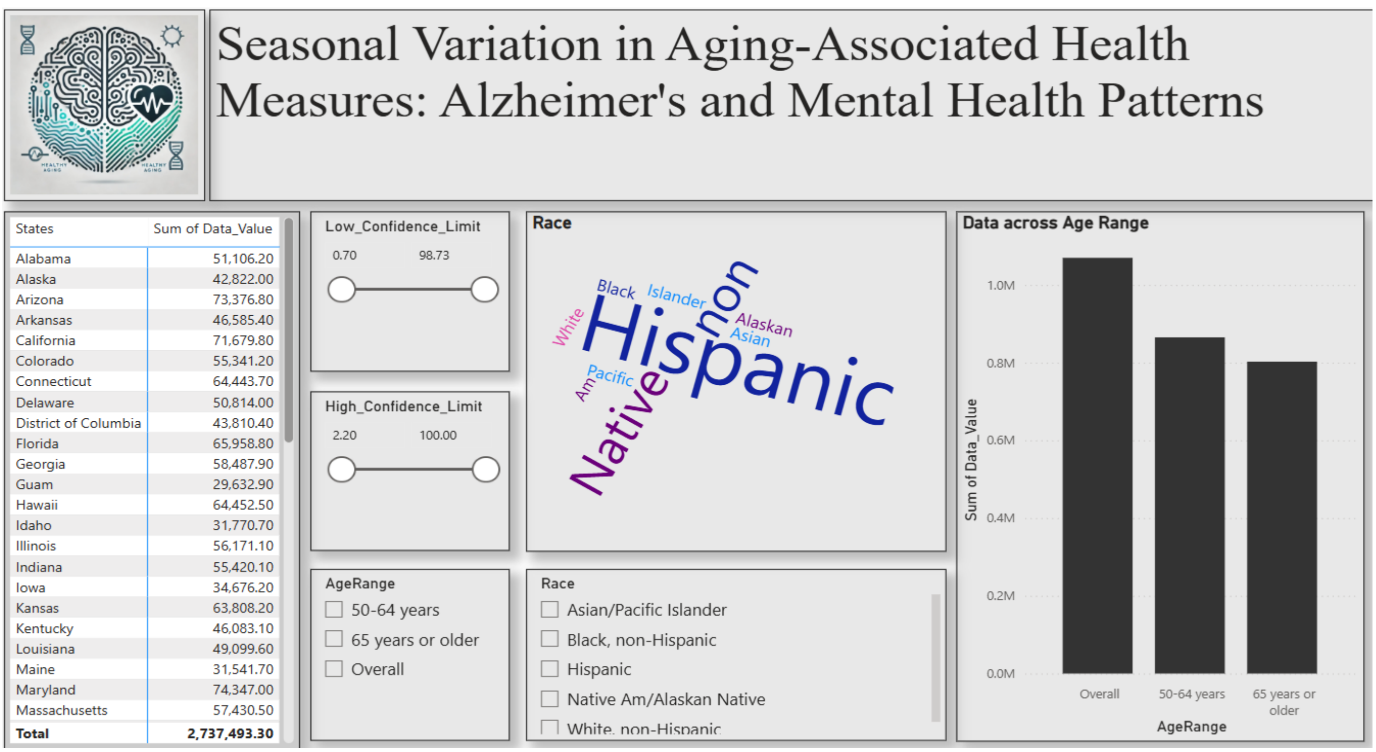


Fig. 3. California, Arizona, and Texas lead in health metrics, with age ranges and racial diversity well-represented

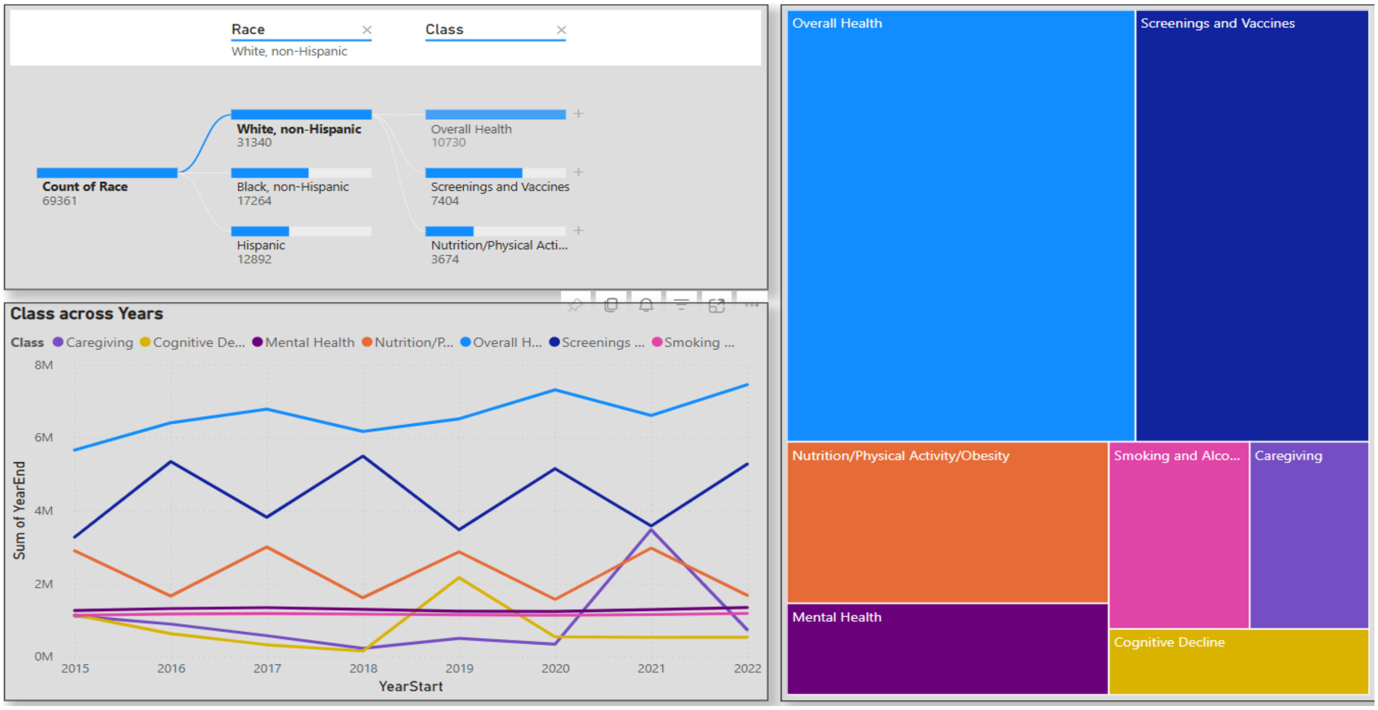
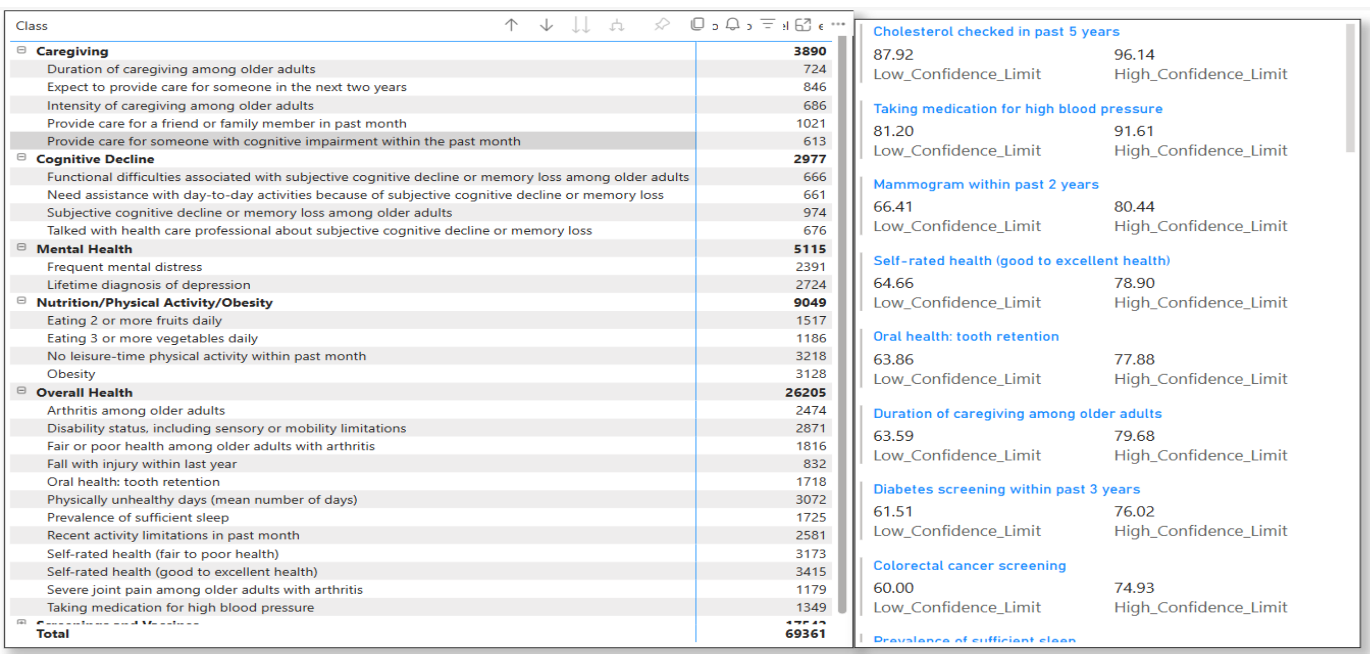


Fig. 4. "Overall Health" dominates across racial groups and years, with a steady increase over time.



Class	
Caregiving	3890
Duration of caregiving among older adults	724
Expect to provide care for someone in the next two years	846
Intensity of caregiving among older adults	686
Provide care for a friend or family member in past month	1021
Provide care for someone with cognitive impairment within the past month	613
Cognitive Decline	2977
Functional difficulties associated with subjective cognitive decline or memory loss among older adults	666
Need assistance with day-to-day activities because of subjective cognitive decline or memory loss	661
Subjective cognitive decline or memory loss among older adults	974
Talked with health care professional about subjective cognitive decline or memory loss	676
Mental Health	5115
Frequent mental distress	2391
Lifetime diagnosis of depression	2724
Nutrition/Physical Activity/Obesity	9049
Eating 2 or more fruits daily	1517
Eating 3 or more vegetables daily	1186
No leisure-time physical activity within past month	3218
Obesity	3128
Overall Health	26205
Arthritis among older adults	2474
Disability status, including sensory or mobility limitations	2871
Fair or poor health among older adults with arthritis	1816
Fall with injury within last year	832
Oral health: tooth retention	1718
Physically unhealthy days (mean number of days)	3072
Prevalence of sufficient sleep	1725
Recent activity limitations in past month	2581
Self-rated health (fair to poor health)	3173
Self-rated health (good to excellent health)	3415
Severe joint pain among older adults with arthritis	1179
Taking medication for high blood pressure	1349
Total	69361

Fig. 5. Preventive health measures like screenings and medications show high prevalence.

VI. Lesson Learned

This project had a number of important lessons in effectively integrating data analysis and visualization in the context of public health. These are elaborated on below as a means of providing full understanding of the challenges faced and solutions devised to address them.

A. Importance of Data Preprocessing and Cleaning

Among the key lessons that were learned was the importance of data preprocessing and cleaning. Most raw datasets contained errors, missing values, or redundancies that could give bias to the results. For instance, cleaning of the Alzheimer's Disease and Healthy Aging dataset was tricky while aligning different fields, keeping demographic and spatial information uniform. Preprocessing included filling gaps in records, normalizing record formats, and removing irrelevant data entries. These efforts were vital in assuring the accuracy and reliability of the visualized trends. Without clean data, even the most advanced visualizations can mislead decision-makers, underlining the necessity of rigorous data preparation for any analytics project.

B. Balancing Granularity and Clarity in Visualizations

Another key takeaway was the challenge of achieving the right balance between granularity and clarity in the visualizations. Highly granular data can provide detailed insights but risks overwhelming stakeholders with complexity. Conversely, overly simplified representations may obscure valuable patterns or nuances. For instance, when visualizing seasonal trends in health metrics, presenting detailed monthly data was more effective for researchers, while aggregated quarterly trends were better suited for policymakers. This experience reinforced the need to tailor visualizations to the target audience while maintaining the integrity of the underlying data.

C. Harnessing Power BI's Interactive Features

The interactive capabilities of Power BI were instrumental in transforming static data into dynamic decision-making tools. These allow stakeholders to take a closer look at data from multiple perspectives: slicers, drill-downs, cross-filtering. The policy analyst would quickly see trends at both state and national levels or zero in on select demographic segments, like rural elderly adults. These functionalities can drive insights that could not be identified on static reports. The takeaway is crystal clear: interactiveness not only makes user experiences better, but it allows the exploration of data to a greater depth and with a much more intuitive understanding of complex health metrics.

D. Collaboration Between Technical and Domain Experts

Collaboration rose as a keystone to success in the project. Expertise from technical experts in data processing, data visualization, and tool development combined with the domain experts on public health metrics. It came in handy, particularly for the interpretation of complex datasets into usable insights. For example, domain experts provided context around statistical anomalies, such as unexpected seasonal spikes in cases of Alzheimer's, while the technical team translated those into user-friendly visualizations. That effort underlined the importance of this interdisciplinary collaboration in filling a gap between data science and deep domain-specific knowledge.

E. Iterative Design and Continuous Feedback

This iterative nature of the design process helped a lot in refining these visualizations. Initial prototypes shared with the stakeholders highlighted areas where they could be improved further-for example, color, simpler legends, or the addition of more filters. All this feedback was brought in, iteration after iteration, which gave a very accessible and impactful set of dashboards. This, at the same time, highlights stakeholder engagement throughout the entire life cycle of the project. Involving end users actively ensures that their needs and preferences are met when the product is finalized, hence a likelihood of increased adoption and utility.

F. Larger Impact on Public Health Analytics

Aside from purely technical lessons learned, this project shed light on broader implications for the emerging field of public health analytics. An integration of these diverse datasets exposed disparities and trends not otherwise apparent through isolated analyses. This reinforced the fact that a holistic approach was needed in health data analytics, an approach that merged several sources of data and perspectives together for a comprehensive view of the public health challenge. This project demonstrated how accessible visualizations can democratize insights from data and thereby allow a wide range of stakeholders, from policy makers to healthcare providers, to make informed decisions.

VII. Future Work

While this project delivered significant insights, it also uncovered avenues for future exploration to enhance its scope and impact. One key area for improvement is integrating real-time data streams into the dashboards, allowing for up-to-date and responsive analysis. Real-time monitoring could significantly improve decision-making during emergencies or seasonal health crises, empowering healthcare professionals to implement timely interventions. Moreover, the integration of predictive analytics can enhance the dashboard to a higher level by forecasting health trends and resource requirements, thereby enhancing preparedness in the health industry.

Another promising direction would be to extend the dataset with international aging health metrics. By analyzing international datasets, comparative studies could identify cultural, environmental, and systemic factors influencing aging health outcomes. Such a perspective would enrich the dashboards, putting into light the holistic view of aging-related health challenges worldwide and encouraging cross-country knowledge sharing.

These advances could have a great impact on the health industry. Real-time and predictive capabilities can help public health agencies move from reactive to proactive strategies, thereby drastically reducing the burden on healthcare systems. With improved vision into global health trends, there would be the possibility of extracting universally best practices to inform better health access and equity globally for an aging population. Future efforts through innovation in data visualization and analytics will ensure health systems remain responsive to the evolving needs of aging populations, realizing healthier societies and more effective public health strategies.

VIII. Benefits on the Industry

The motivation for this work came from the increasing burden of AD and the need for effective analysis and visualization methods for data associated with public health. Alzheimer's Disease is one of the fastest-growing concerns for millions of people across the globe, and there is an urgent need for new ideas that can help in early diagnosis, treatment, and management. The motivation for this project seems to come from the urge to leverage data-driven approaches in uncovering trends, patterns, and insights that may inform public health efforts, healthcare professionals, and policy makers. By integrating data analysis and visualization, we seek to provide comprehensive understanding of the disease's progression and associated risk factors that will eventually contribute toward better healthcare outcomes among affected individuals and their families.

From the perspective of industry, this project has done something of great impact and consequence to the health and pharmaceutical world. These will enable much better analysis of large data sets and better visualization of complex information, which inform decision-making at all levels in healthcare. For instance, predictive analytics of Alzheimer's Disease can facilitate an early diagnosis whereby health professionals can intervene earlier. This might slow down further degeneration and allow the patients a better quality of life. Moreover, insights from visualization can help pharmaceutical companies in observing drug efficacy trends or how different treatments affect patient outcomes. With such data-driven strategies in place, the healthcare industry may adopt proactive steps in managing AD and eventually reduce long-term healthcare costs and increase the overall standard of care.

Apart from the field of health, the project holds huge potential for further influence in public health policymaking and drives innovation across all industries. The use of advanced data analytics and visualization tools provides valuable insights that can guide the allocation of resources, the design of public health campaigns, and the development of new healthcare technologies. The results from this study would add to the ongoing efforts towards shaping a more responsive, efficient, and future-prepared industry in dealing with public health challenges amid aging populations and chronic diseases such as Alzheimer's. Moreover, it shows the transformational impact that data-driven solutions can have on many of the most challenging societal problems and will undoubtedly encourage similar approaches in many other areas, including social services, insurance, and urban planning. Addressing Alzheimer's Disease with such innovative methodologies has the potential to influence not only individual patient outcomes but also industry-wide practices and policies that can improve the health of whole populations.

IX. Conclusion

This study successfully demonstrated the transformative potential of interactive data visualization in the analysis and addressing of health metrics related to aging. Using the Alzheimer's Disease and Healthy Ageing dataset, we were able to identify key demographic disparities, spatial patterns, and seasonal variations that are driving public health outcomes. The results of this study show how aging populations are being disproportionately affected by environmental and socioeconomic factors, hence requiring targeted interventions. With integrated Power BI dashboards, policymakers have access to an intuitive and interactive tool that informs strategies on reducing inequities in health care access and outcomes.

The project also brought into light the power of visual storytelling in driving actionable insights. The intuitive visualization of data presented on the dashboards created a seamless bridge between complex datasets and user-friendly analysis, fostering a deeper understanding among healthcare professionals and decision-makers. This approach not only streamlines public health planning but also ensures that interventions are data-driven and effectively address the nuanced needs of aging populations.

The health industry will benefit a great deal from such innovations. Policymakers and healthcare providers can use these tools to monitor aging trends in real time, identify vulnerable regions or groups, and deploy resources more efficiently. This project also highlighted the importance of accessible and transparent data visualization tools in fostering collaboration across stakeholders-from researchers to healthcare administrators-for better health outcomes.

Acknowledgment

We express our sincere gratitude to all individuals and organizations who contributed to the success of this project. We are particularly grateful to the U.S. Department of Health & Human Services for providing the Alzheimer's Disease and Healthy Ageing dataset, which served as the foundation for our analysis. We extend our heartfelt thanks to Professor Shih Yu Chang and Teaching Assistant Sourab Rajendra Saklecha at San Jose State University for their invaluable guidance, encouragement, and insightful feedback throughout the project. Their expertise and mentorship were instrumental in shaping our approach and deepening our understanding of the subject. We also thank San Jose State University for providing a supportive and resourceful academic environment that fostered the growth and development of this work. Special thanks go to our peers and colleagues for their constructive critiques and shared expertise, which enriched our efforts. Finally, we acknowledge the creators and developers of Power BI, whose tool enabled the transformation of complex data into meaningful and actionable insights. This project would not have been possible without the collective efforts and support of all those involved.

References

[1] C. Birkenbihl, Y. Salimi, D. Domingo-Fernández, S. Lovestone, AddNeuroMed Consortium, H. Fröhlich, et al., "Evaluating the Alzheimer's disease data landscape," *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, vol. 6, no. 1, pp. e12102, 2020.

[2] K. Seo, R. Pan, D. Lee, P. Thiyyagura, and K. Chen, "Visualizing Alzheimer's disease progression in low dimensional manifolds," *Heliyon*, vol. 5, no. 8, 2019.

[3] R. Guan, X. Wen, Y. Liang, D. Xu, B. He, and X. Feng, "Trends in Alzheimer's disease research based upon machine learning analysis of PubMed abstracts," *International Journal of Biological Sciences*, vol. 15, no. 10, pp. 2065, 2019.

[4] M. Song, G. E. Heo, and D. Lee, "Identifying the landscape of Alzheimer's disease research with network and content analysis," *Scientometrics*, vol. 102, pp. 905–927, 2015.

[5] D. D. Martinelli, "Evolution of Alzheimer's disease research from a health-tech perspective: Insights from text mining," *International Journal of Information Management Data Insights*, vol. 2, no. 2, pp. 100089, 2022.

[6] A. Malvaso, A. Gatti, G. Negro, C. Calatozzolo, V. Medici, and T. E. Poloni, "Microglial senescence and activation in healthy aging and Alzheimer's disease: systematic review and neuropathological scoring," *Cells*, vol. 12, no. 24, pp. 2824, 2023.

[7] M. N. Kane, "Mental health issues and Alzheimer's disease," *American Journal of Alzheimer's Disease*, vol. 14, no. 2, pp. 102–110, 1999.

[8] L. Li, Y. Huang, Y. Han, and J. Jiang, "Use of deep learning genomics to discriminate Alzheimer's disease and healthy controls," in *Proc. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, San Diego, CA, USA, 2021, pp. 5788–5791.

[9] N. A. Bihlmeyer, E. Merrill, Y. Lambert, G. P. Srivastava, T. W. Clark, B. T. Hyman, and S. Das, "Novel methods for integration and visualization of genomics and genetics data in Alzheimer's disease," *Alzheimer's & Dementia*, vol. 15, no. 6, pp. 788–798, 2019.