COPY RIGHT

ELSEVIER SSRN

Paper Authors

Dr. Ch. Rajendra Babu, Anusha Nandam, Reethika Kolusu, Nagendra Sureddi

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

Childhood Allergies Analysis

Dr. Ch. Rajendra Babu¹ - Professor & HOD, Department of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India.
Anusha Nandam², Department of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India. (19HP1A0567).
Reethika Kolusu³, Department of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India. (19HP1A0583).
Nagendra Sureddi⁴, Department of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India. (19HP1A05B2).

Abstract
According to the Centers for Disease Control and Prevention (CDC), food allergies in children have risen by 50% since 1990. That is one in every 13 children now has food allergies. Hence there is a need to deeply study different types of allergies based on various factors like gender, race, and age, but studying large datasets with a huge number of entries written in text format/table format is difficult. Hence there is a need to build an Analytical solution that processes a regularly updating dataset from the web using technologies like Azure Data Factory and Snowflake(Snowpark) for visualization in Power BI for a better and quick understanding of data which leads to early preventive measures of childhood allergies.

Keywords: Azure, Snowflake, Power BI.

I. Introduction

Food allergies are a growing concern in childhood health, affecting millions of children worldwide. A food allergy is an adverse immune response to a particular food protein that can result in a range of symptoms, from mild to life-threatening. The prevalence of childhood food allergies has increased significantly in recent decades, with estimates suggesting that up to 8% of children worldwide have a food allergy.

The causes of childhood food allergies are not entirely understood, but research suggests that a combination of genetic and environmental factors may play a role. Common food allergens in children include cow's milk, eggs, peanuts, tree nuts, fish, shellfish, soy, and wheat. Symptoms of a food allergy can vary from mild to severe and may include hives, swelling, vomiting, and difficulty breathing.

The purpose of this analysis project is to provide a comprehensive overview of the child using Power BI dashboards and Reports.

Our dataset consists of 50 columns and 3,30,000 rows approximately, which provides the status of 19 different types of allergies at the start of the survey and end of the survey of children along with categorical columns like gender, age, race, ethnicity, payer factors, etc.

II. Literature Survey

"The Natural History of Childhood Asthma: Insights from the Tucson Cohort Study" by Fernando D. Martinez et al. (1995) - This paper is a comprehensive review of the natural history of childhood asthma, including risk factors, clinical manifestations, and long-term outcomes.
"Prevalence of Childhood Food Allergy in the United States" by Ruchi S. Gupta et al. (2011) - This paper is a landmark study that estimated the prevalence of food allergy among children in the United States. The study found that approximately 8% of children had a food allergy.

"The Global Allergy Epidemic: The Role of Industrialization" by Martin J. Blaser et al. (2014) - This paper examines the role

of industrialization in the global allergy epidemic, which has seen a significant increase in the prevalence of allergic diseases in recent decades.

"Early-Life Environmental Exposures and Childhood Asthma" by Susan L. Prescott et al. (2013) - This paper discusses the relationship between early-life environmental exposures and the development of childhood asthma, including factors such as diet, microbiome, and exposure to allergens and pollutants.

"The Atopic March: Progression from Atopic Dermatitis to Allergic Rhinitis and Asthma" by Donald Y. M. Leung et al. (2016) - This paper reviews the evidence for the "atopic march," which refers to the progression of allergic diseases from atopic dermatitis in infancy to allergic rhinitis and asthma in later childhood and adolescence.

III. Methodology

The proposed system is divided into four modules based on four technologies namely Azure, Snowflake, Power BI, and DevOps.
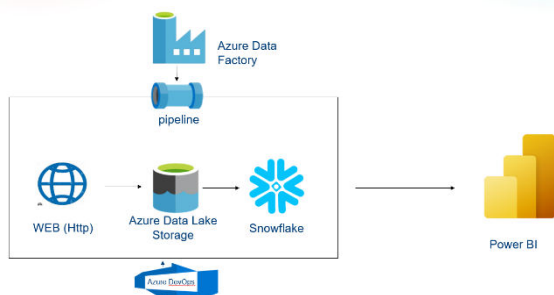
Azure is a cloud computing platform and an online portal that allows you to access and manage cloud services and resources provided by Microsoft. We are going to use three resources ADLS, ADF, and Logic App.

We use ADLS(Azure Data lake storage gen 2) to store our dataset and then created two logic apps for email alert and power BI refreshing, and finally created ADF(Azure data factory) for automating the total workflow.
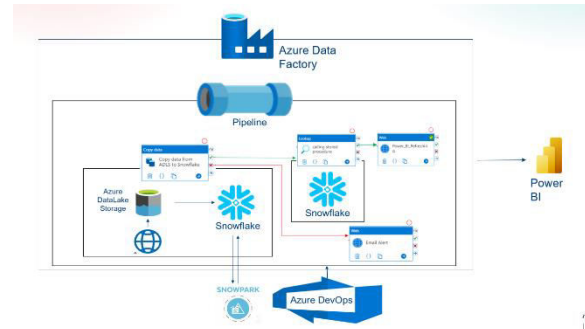

In ADF we have created a pipeline that automates the total workflow which consists of three activities (copy data activity, look up activity, and web activities).
We first copy data from the web to ADLS then to Snowflake using copy data activity then use a lookup activity to call the stored procedure and two web activities for email alert and power bi-refreshing.
The dataset copied to snowflake is then transformed into a curated dataset with the help of stored procedure created in snowflake with the help of snowpark library which allows us to use python and its libraries like pandas in the snowlight interface other than Sequel which is the only language supported by Snowflake.
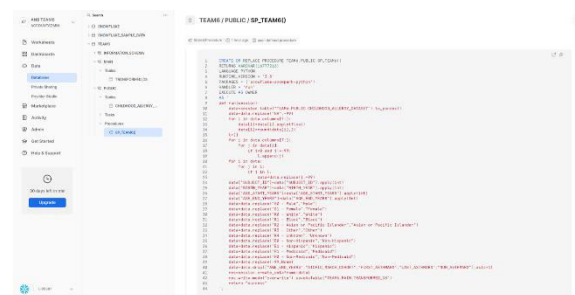
After transformations we connect power BI to our snowflake account using get data.
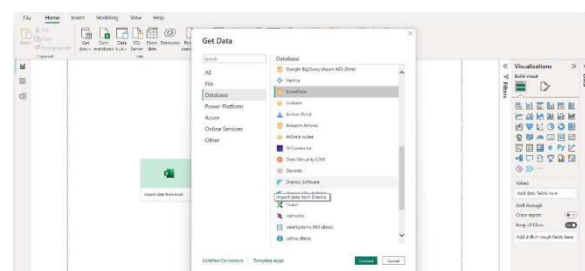
In power BI first, we have to write a DAX measure to make our data convenient for analysis, then we use different types of visualizations like clustered column charts, funnel charts, pie charts, slicers ,etc. for representing our dataset for clear and faster understanding.
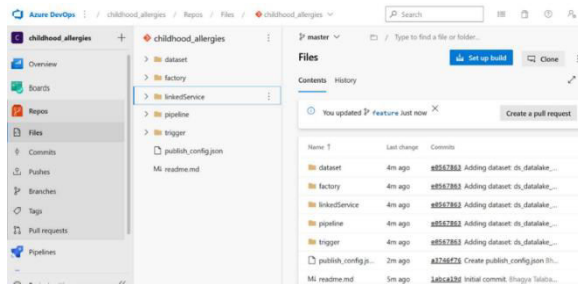
Then we connect to DevOps for
·Collaboration, version control, and Better performance.

IV.   Results and Analysis
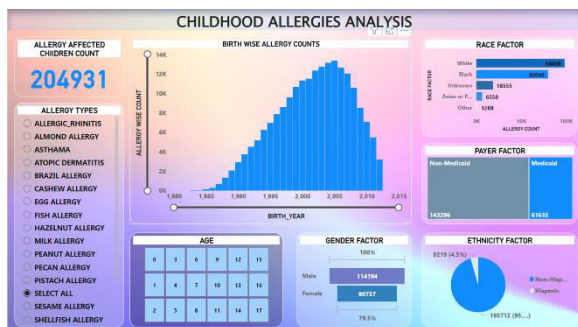
There are a total of 204931 allergy-affected children present in our dataset.
Most of the children are affected by Asthma (63874)
Male children (114194) are comparatively more affected than Female children (90737)
Children who were born in the year 2005 got affected highly (13375)
Approximately 49% of affected children are whites. (Race)
Children who were born in the year 1983 are less prone to allergies (8)

Children who were affected by allergies are mostly not supported by the government (Non –Medicaid)
Only one child got affected by Tree Nut Allergy.
Children who were just born (Age-0) in the year 2005 highly got affected (10823)
Children who were born in the first decade of the 21st Century (2000-2010) were highly affected [31471] by Atopic Dermatitis.

V.   Conclusion

We conclude that giving scientists and doctor's access to a power-bi visualization dashboard will enable them to study kid allergies more swiftly and conveniently.

VI.   Future scope

Future analysis of childhood allergy datasets has a lot of potential. Analysing the trends, causes, and treatment options will aid in establishing better allergy- preventative measures and therapies due to the rising prevalence of allergies in children.

VII.   References

[1]https://ieeexplore.ieee.org/document/7807236
[2]https://www.sciencedirect.com/science/article/abs/pii/S1526054202001884
[3]https://www.sciencedirect.com/science/article/abs/pii/S1081120617307299
[4]https://www.thelancet.com/journals/lanchi/article/PIIS2352-4642(22)00215-2/fulltext
[5]https://onlinelibrary.wiley.com/doi/10.1111/pai.13262
[6]https://pubmed.ncbi.nlm.nih.gov/35985346/