Factors Leading to Asthma Diagnosis

Author

Gabriel Kanife

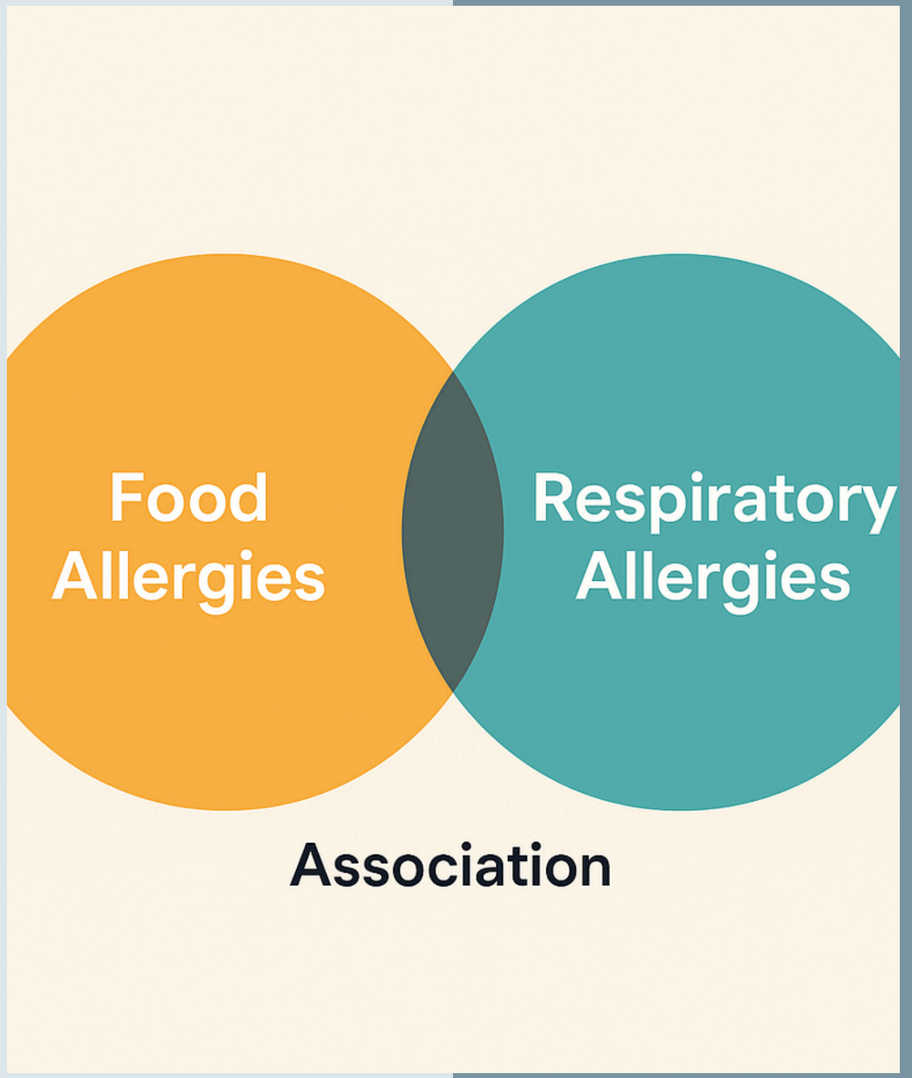
My name is Gabriel Kanife, and I’m a Senior majoring in Data Intelligence and Business Analytics at the University of West Georgia.

Over the past semester, I’ve been building my skills in data analysis using R, and this project has given me the chance to apply those skills to a real-world dataset. I’m especially interested in using data to uncover patterns and support better decision-making in health and business environments.

This presentation reflects my progress in learning how to clean, analyze, and interpret data — and I’m excited to share what I found.

Why?

The rates of childhood allergic conditions are changing, prompting the need for continued surveillance. Studies have shown a Medical Connection between Food Allergies and Respiratory Allergies.



The Data

The epidemiologic characteristics of healthcare provider-diagnosed eczema, asthma, allergic rhinitis, and food allergy in children: a retrospective cohort study

https://bmcpediatr.biomedcentral.com/articles/10.1186/s12887-016-0673-z

	SUBJECT_ID	BIRTH_YEAR	GENDER_FACTOR	RACE_FACTOR	ETHNICITY_FACTOR	PAYER_FACTOR	ATOPIC_MARCH_COHORT	AGE_START_YEARS	AGE_END_YEARS
4	4	2004	S0 – Male	R4 – Unknown	E1 – Hispanic	P0 – Non-Medicaid	FALSE	2.398357290	9.111567
5	5	2006	S1 – Female	R1 – Black	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.013689254	6.193018
6	6	2006	S0 – Male	R0 – White	E0 – Non-Hispanic	P1 – Medicaid	FALSE	0.087611225	4.958248
7	7	2006	S1 – Female	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.010951403	6.677618
8	8	2006	S1 – Female	R1 – Black	E0 – Non-Hispanic	P0 – Non-Medicaid	TRUE	0.032854209	6.286105
9	9	2006	S0 – Male	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	TRUE	0.019164956	6.726899
10	10	2006	S1 – Female	R1 – Black	E0 – Non-Hispanic	P1 – Medicaid	TRUE	0.021902806	6.275154
11	11	2006	S0 – Male	R4 – Unknown	E0 – Non-Hispanic	P0 – Non-Medicaid	TRUE	0.019164956	7.003422
12	12	2005	S0 – Male	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	1.897330595	5.900068
13	13	2006	S1 – Female	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	TRUE	0.016427105	6.663929
14	14	2006	S0 – Male	R4 – Unknown	E1 – Hispanic	P1 – Medicaid	FALSE	0.013689254	5.911020
15	15	2006	S1 – Female	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.013689254	2.866530
16	16	2006	S1 – Female	R1 – Black	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.024640657	1.226557
17	17	2006	S0 – Male	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	TRUE	0.010951403	6.217659
18	18	2006	S0 – Male	R4 – Unknown	E0 – Non-Hispanic	P1 – Medicaid	FALSE	0.013689254	1.097878
19	19	2006	S0 – Male	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	TRUE	0.016427105	6.187543
20	20	2006	S0 – Male	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.013689254	2.841889
21	21	2006	S1 – Female	R1 – Black	E0 – Non-Hispanic	P1 – Medicaid	FALSE	0.013689254	7.049966
22	22	2006	S1 – Female	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.013689254	6.428474
23	23	2003	S1 – Female	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	3.288158795	10.160164
24	24	1999	S1 – Female	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	7.783709788	14.704997
25	25	2006	S1 – Female	R4 – Unknown	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.016427105	6.220397
26	26	2006	S1 – Female	R0 – White	E1 – Hispanic	P0 – Non-Medicaid	TRUE	0.010951403	7.047228
27	27	2006	S1 – Female	R4 – Unknown	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	0.071184120	1.289528
28	28	1999	S1 – Female	R0 – White	E0 – Non-Hispanic	P0 – Non-Medicaid	FALSE	7.581108830	14.398357
29	29	2006	S0 – Male	R2 – Asian or Pacific Islander	E0 – Non-Hispanic	P0 – Non-Medicaid	TRUE	0.019164956	6.584531
30	30	2006	S1 – Female	R1 – Black	E0 – Non-Hispanic	P1 – Medicaid	FALSE	2.507871321	7.011636
31	31	2006	S0 – Male	R1 – Black	E0 – Non-Hispanic	P1 – Medicaid	TRUE	0.021902806	5.681040

Interesting Facts about the data.

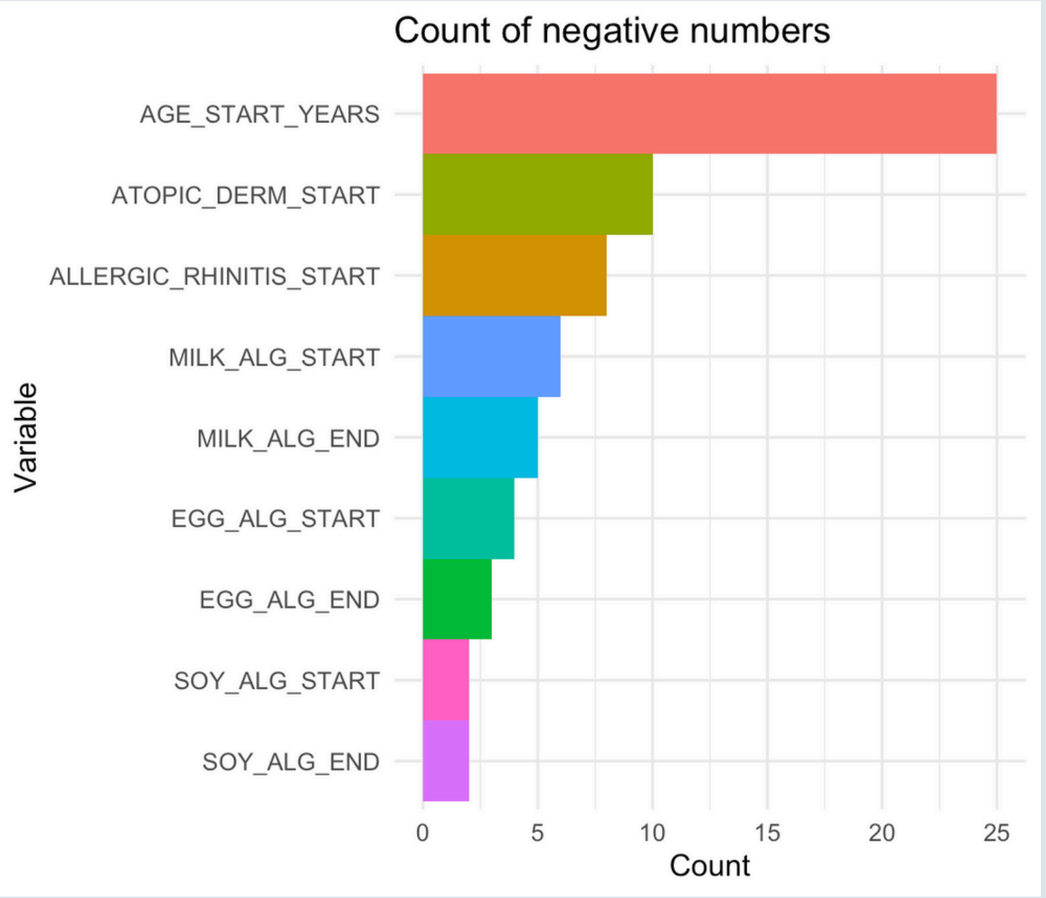
- For some kids, the study began before they were born.
- Some patients were admistered Asthma medications without having the allergy.

Data Cleaning

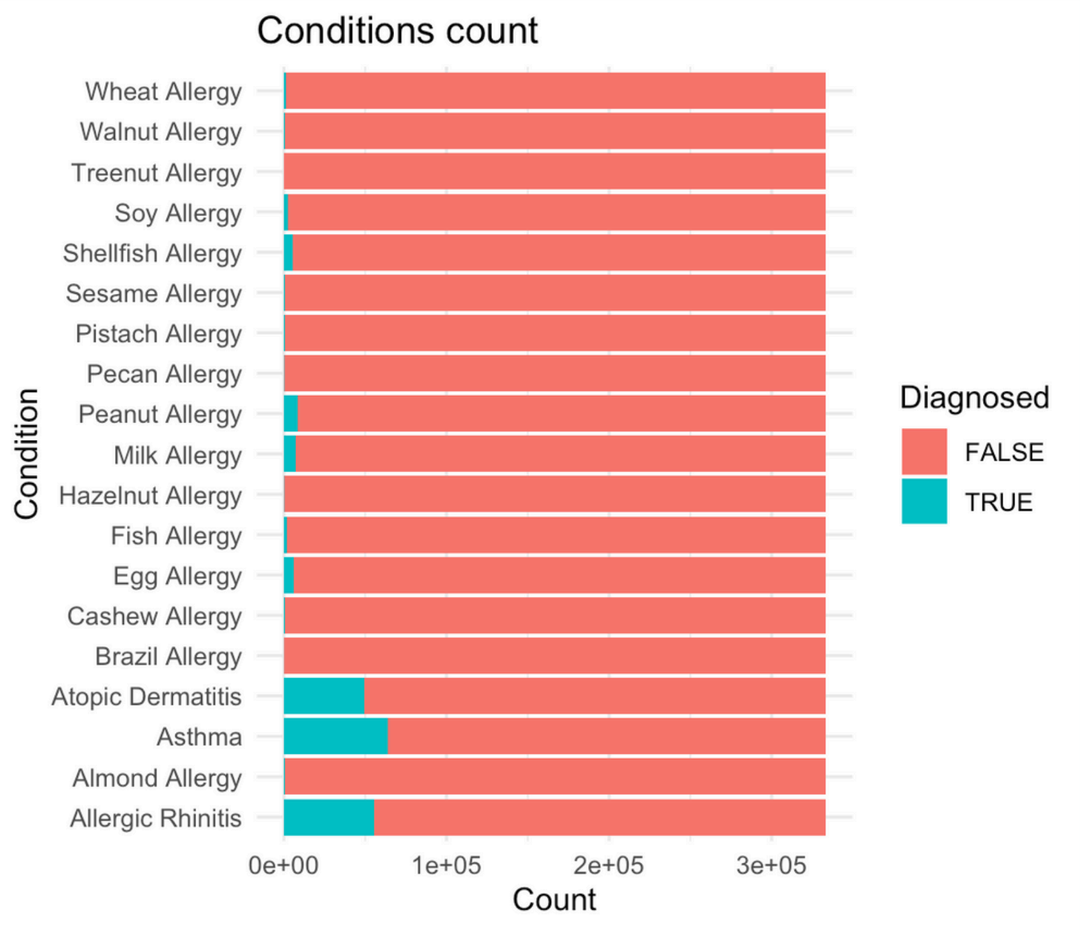
The first step in cleaning this data was to check for negative numbers. As this can affect the analysis, all negative numbers were removed.

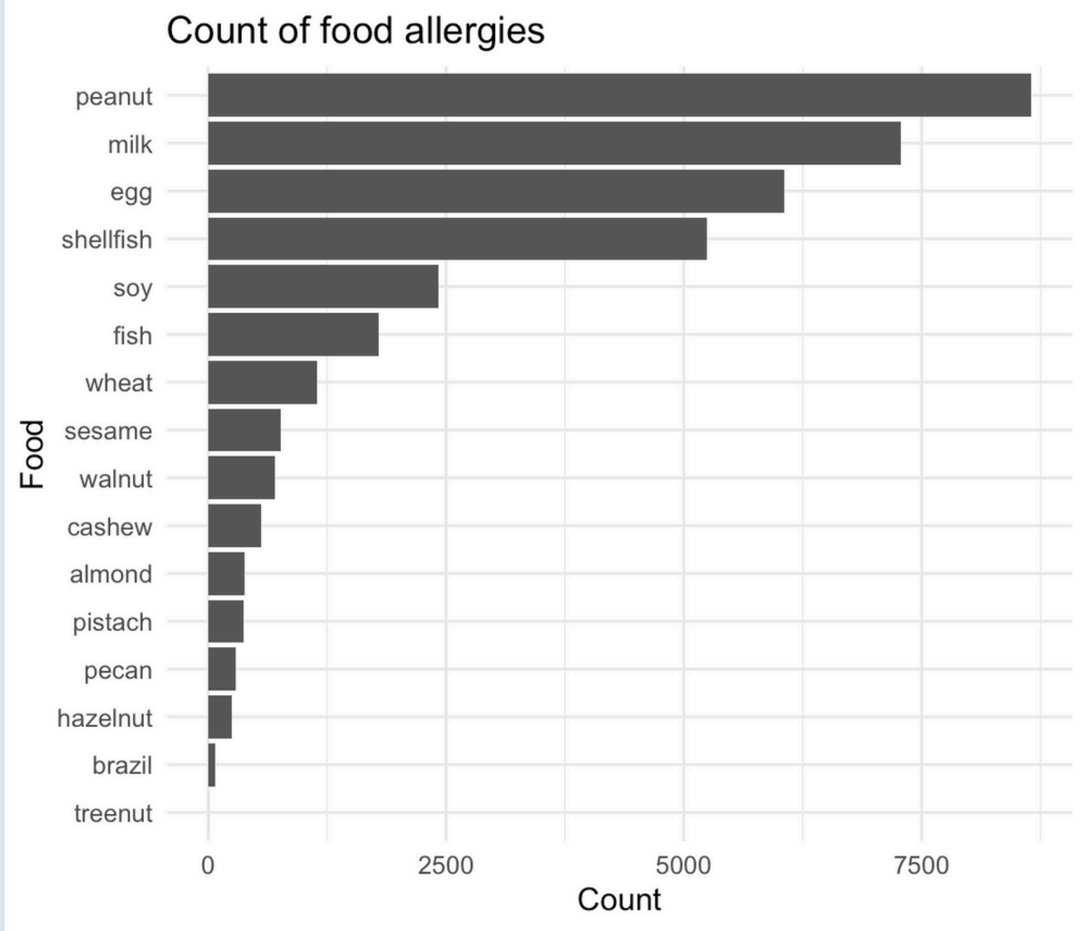
<u>Out Come</u>

- A total OF 53 patients have values that are less than 0
- Number of patients before exclusion: 333200
- Number of patients after exclusion: 333147

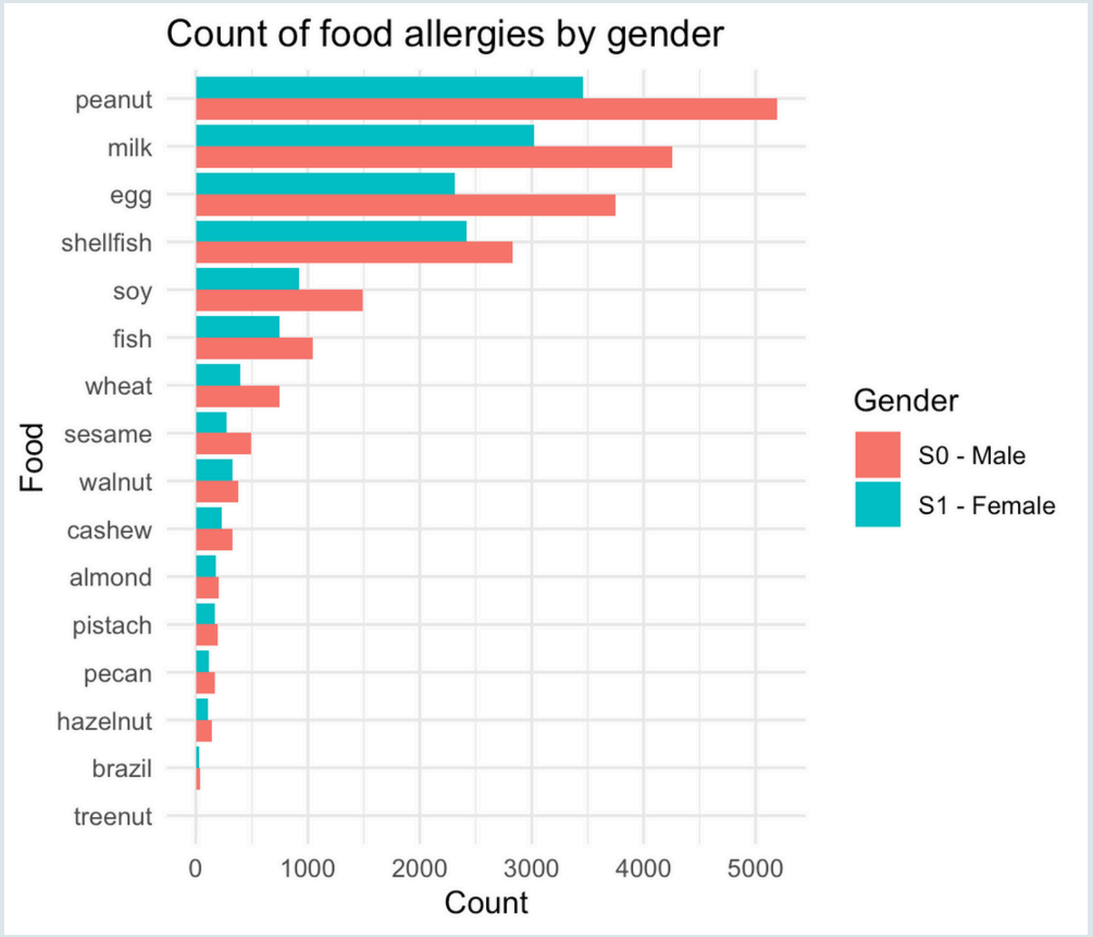


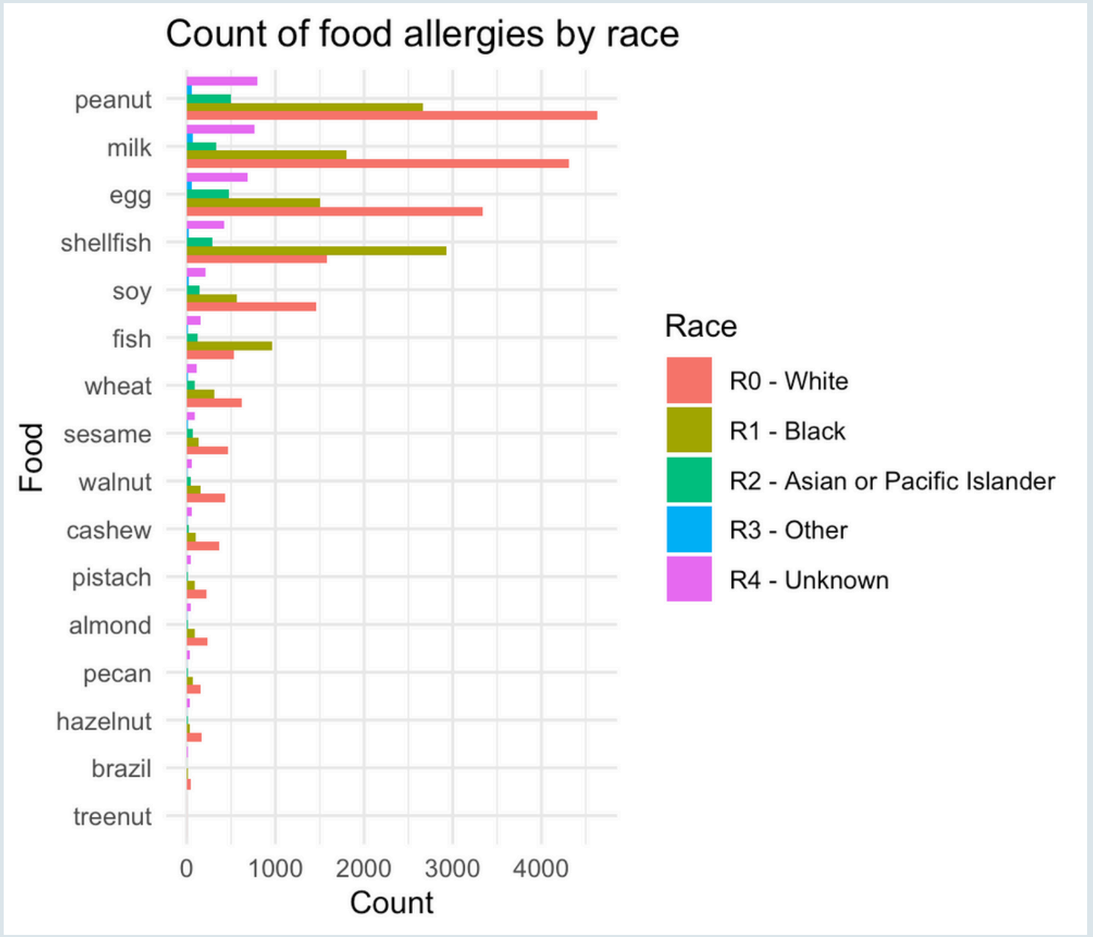
Data Attributes



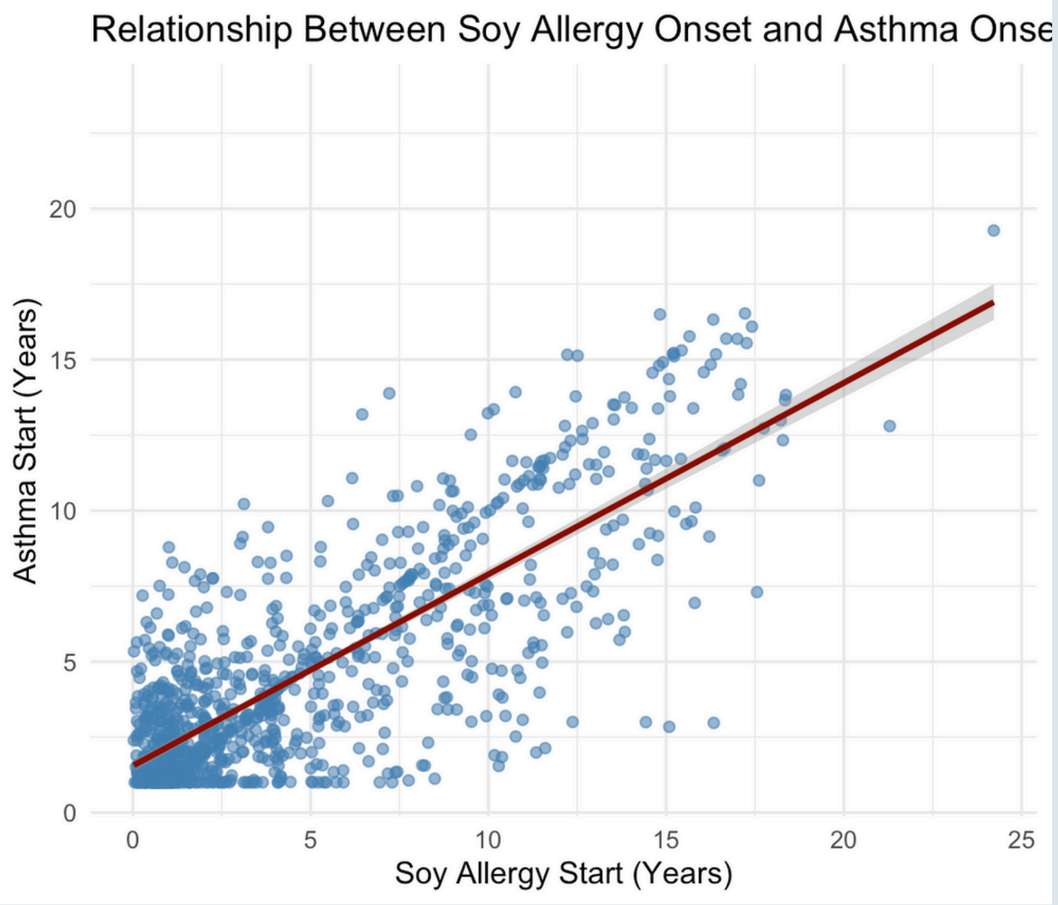


Data Attributes





Data Analysis
Correlation Analysis

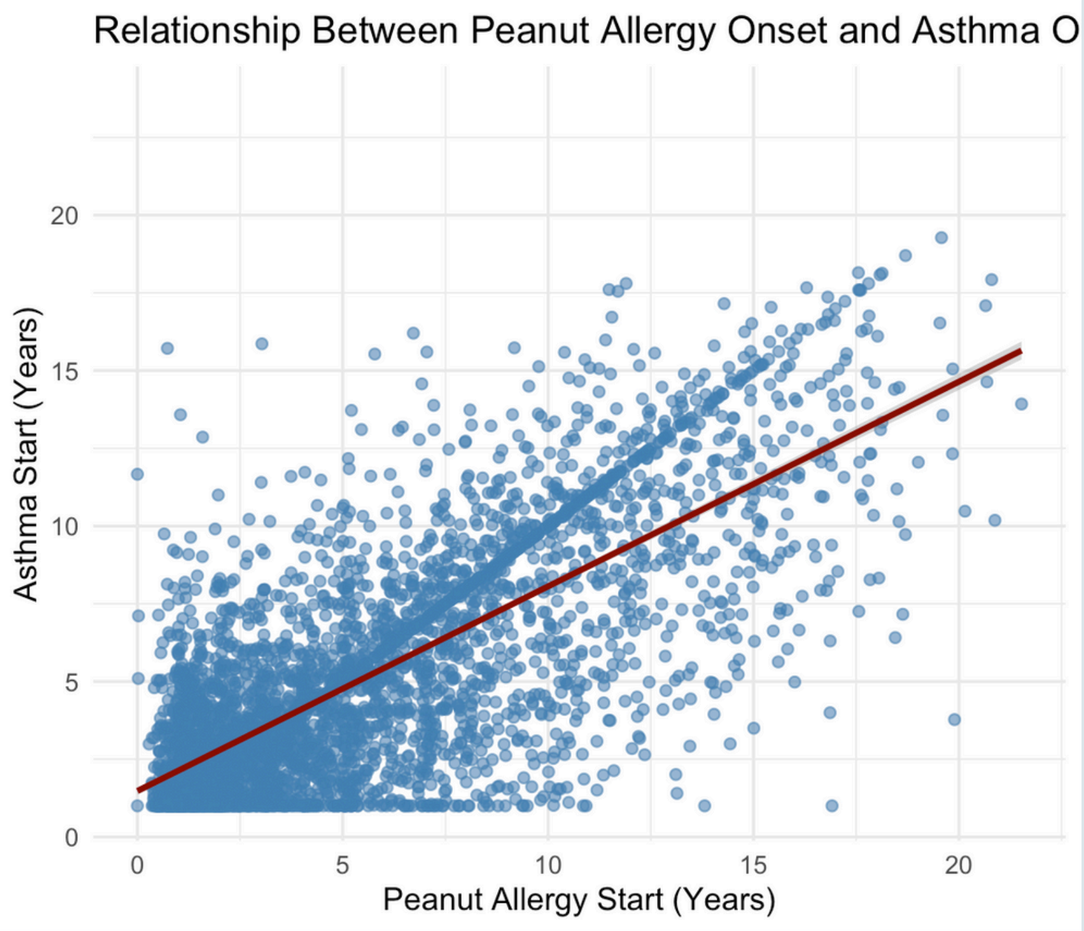


There is a strong positive linear correlation (about 0.80) between the onset of soy allergy and asthma.

As the age of soy allergy onset increases, asthma onset tends to increase as well (i.e., they occur around the same time).

Data Analysis

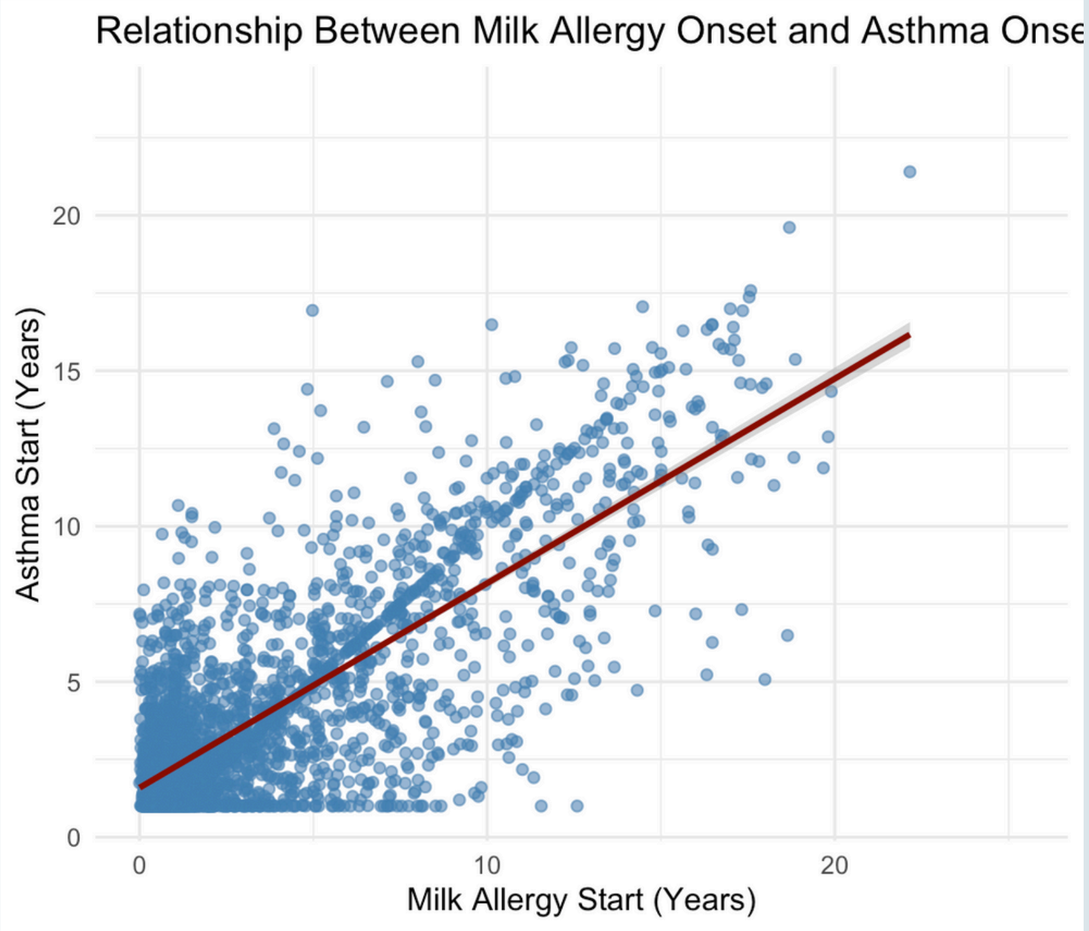
Correlation Analysis



There is a strong positive linear correlation (about 0.80) between the onset of Peanut allergy and asthma.

As the age of peanut allergy onset increases, asthma onset tends to increase as well (i.e., they occur around the same time).

Data Analysis
Correlation Analysis



There is a strong positive linear correlation (about 0.77) between the onset of Milk allergy and asthma.

This represents a strong positive linear relationship, indicating that an earlier onset of milk allergy tends to be associated with an earlier onset of asthma.

Data Analysis
Regression Analysis

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
2	(Intercept)	1.518 ***	1.694 ***	0.690 *	-0.037	-0.041	1.720 ***
3		(0.106)	(0.150)	(0.357)	(1.297)	(1.331)	(0.116)
4	Milk Allergy Start	0.385 ***					0.403 ***
5		(0.061)					(0.060)
6	Soy Allergy Start	0.238 ***					0.222 ***
7		(0.061)					(0.060)
8	Egg Allergy Start		0.215 ***				
9			(0.064)				
10	Wheat Allergy Start		0.395 ***				
11			(0.065)				
12	Peanut Allergy Start			0.324 ***			
13				(0.091)			
14	Walnut Allergy Start			0.335 ***			
15				(0.086)			
16	Almond Allergy Start				-0.031	-0.031	
17					(0.940)	(0.962)	
18	Hazelnut Allergy Start				0.640	0.639	
19					(0.955)	(0.978)	
20	as.factor(payer_factor)P1 -					0.082	-0.706 ***
21						(1.740)	(0.174)
22	N	516	368	141	25	25	516
23	R2	0.541	0.517	0.563	0.517	0.517	0.555
24	logLik	-1036.997	-787.633	-317.006	-61.967	-61.965	-1028.798
25	AIC	2081.994	1583.266	642.011	131.934	133.931	2067.597
26	*** p < 0.01; ** p < 0.05; * p < 0.1.						
27							

Conclusion

Finally, the findings indicate that the presence of food allergy is a risk factor for the subsequent development of respiratory allergy (Asthma). These findings allow new insights into the epidemiologic characteristics of these diseases, describe the importance of utilizing provider-diagnosis data to complement participant reporting methodologies, and provide important information to shape future efforts aimed at prevention, diagnosis, and management of these common pediatric conditions.

Thank you