

Seasonality and Covid-19 Impacts to the National School Lunch Program

Scout Leonard

Contents

My Question:	1
Background	2
Data Description	2
Data Quality	3
Analysis Plan	3
Seasonality in NSLP	3
Impact of Covid on NSLP	3
Summarize results visually and in words	3
Seasonality of NSLP	4
Impact of Covid on NSLP	5
Next steps and future directions	6
References	6

This statistical analysis was completed for my Master's program course, Environmental Data Science 222: Statistics for Environmental Data Science. Comments or suggestions about the analysis and the complete history of the code can be accessed at my **project repository on Github**. The code is not included in the PDF documentation of the project, but is visible at this link to the live project web page.

My Question:

My analysis seeks to explore the questions: *Are there seasonal differences in how U.S. school food programs feed students? What about long-term trends, and how did any trends change in 2020 during school closures from the Covid-19 pandemic?*

I have previous experience working in school food systems to expand access to this resource for school communities, but you need no time in cafeterias to guess that there would be seasonality in feeding kids. Students off for summer means fewer students participating in the school lunch in July and August. I was more interested in understanding seasonal differences within the school year: for example, *does the US feed more students in fall compared to winter?*

I am also interested in the impact of Covid-19 on school lunch. Across the U.S., many school nutrition programs adjusted to serving grab-and-go meals for students who rely on school for part of their daily nutrition during school closures. Understanding service gaps during this time

Background

The National School Lunch Program (NSLP) is an enormous food system with major implications for equity in American K-12 education systems. Every day, NSLP provides ~30 million children school lunch at free or reduced prices (1). It operates in public and nonprofit private schools and residential childcare facilities (2).

To provide meals at free and reduced cost to students, participating school districts are reimbursed cash subsidies for every qualifying meal they serve. To qualify for subsidy, meals served by Nutrition Services operators must meet federal meal pattern policies which define meal content around qualifying food group combinations, sugar content, etc.

In the US, an estimated 11.7 million children are estimated to live in food insecure homes this year (5). Students who experience hunger are reported to experience higher risk of decreased academic performance, more tardy and absent days, and more mental health challenges (4). A robust understanding of school food systems can create data-informed decisions around USDA and state school food policy and on-the-ground management so that students do not have hunger as an additional barrier to their wellbeing and academic achievement.

Data Description

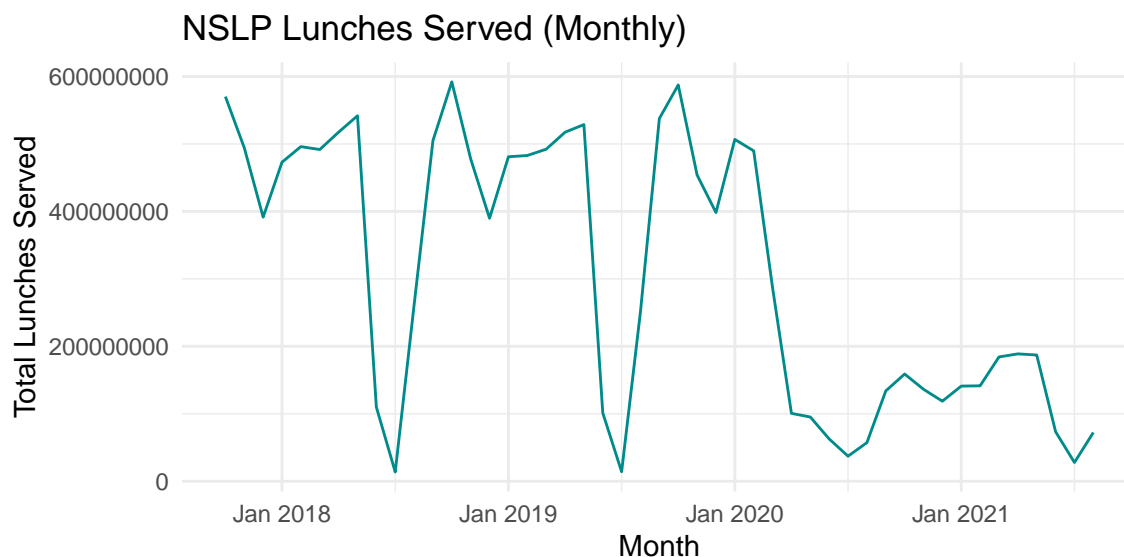
The Food and Nutrition Services sector of the USDA offers monthly and annual reports of national participation in the National School Lunch Program and other school meal programs subsidized by the USDA (breakfast, seamless summer, supper, and snacks).

To answer my questions, I used the National Level Monthly Data for the National School Lunch Program, which gives national-level monthly participation in USDA school nutrition programs since October 2017 (47 months). This data comes in publicly available .pdf and .xlsx formats. I downloaded the .xlsx format of the monthly data and did some tidying in Excel to make sure the .csv version would be friendly with R. This version of the data can be found in my Github repository.

Even after this manipulation in excel, there were changes to be made for easier manipulation of my monthly lunch data dataframe. Here, I give columns more descriptive names, remove odd columns that were added to the dataframe when I read in the .csv, mutate columns to exclude “%” “-” or spaces that R cannot make sense of.

Last, I make sure the class of each column is correct. All of the columns were of class `numeric` when read in, so I mutated the `month` column to `yearmon` using the `zoo` package, and everything else to `numeric`.

Finally, to have an initial look at my data, I make simple line graphs here:



At a first glance, I can see that there are seasonal dips in NSLP lunch participation in the summers, as expected. There also appears to be a dip in participation before January, which I would suspect is due to holiday recesses. It seems that the start of the school year does have higher participation than any of the winter and spring months, though.

I can see that participation dropped dramatically when schools closed, by about 2/3 for all students, and by 1/2 for just the students who qualify for free lunch.

Data Quality

Every school day, school cafeteria staff at participating NSLP schools report meals claimed by students who qualify for free- and reduced-priced lunches. That meets the numbers from this dataset are self-reported by schools and districts nationally. The USDA conducts periodically audits schools to control for accurate serving, counting, and claiming or reimbursable meals.

This analysis assumes that the numbers presented by the USDA are accurate monthly meals served nationally, but there is risk for systematically biased sampling. In my own work, I have observed inaccurate claim reports as a results of short staffing and reduced bandwidth for serving *and* counting. This may mean that school nutrition programs with lower budgets (i.e. more free-and-reduced eligible students, because of funding models) are at risk for less accurate reporting.

Analysis Plan

Seasonality in NSLP

For the first part of my analysis, I investigate seasonal differences in NSLP participation by decomposing the data seasonal and trend components using the `classical_decomposition()` function and applying `autoplot()` to that. The classical decomposition decomposes my data into the following relationship

$$Lt = St + Tt + Rt$$

Where Lt is lunches served per month, St is a seasonal component, and Tt is the trend component.

To visualize this decomposition, I first create a tsibble with the months as class `yearmonth` and the total lunches served per month, then create an additive classical decomposition model. I then use this model to generate an autoplot which helps to visualize the presence of seasonality and long-term trends in the data for both groups. Finally, I generate an autocorrelation function with a lag of 12 because I want to see how much participation in one month is correlated with participation for the rest of the school year.

Impact of Covid on NSLP

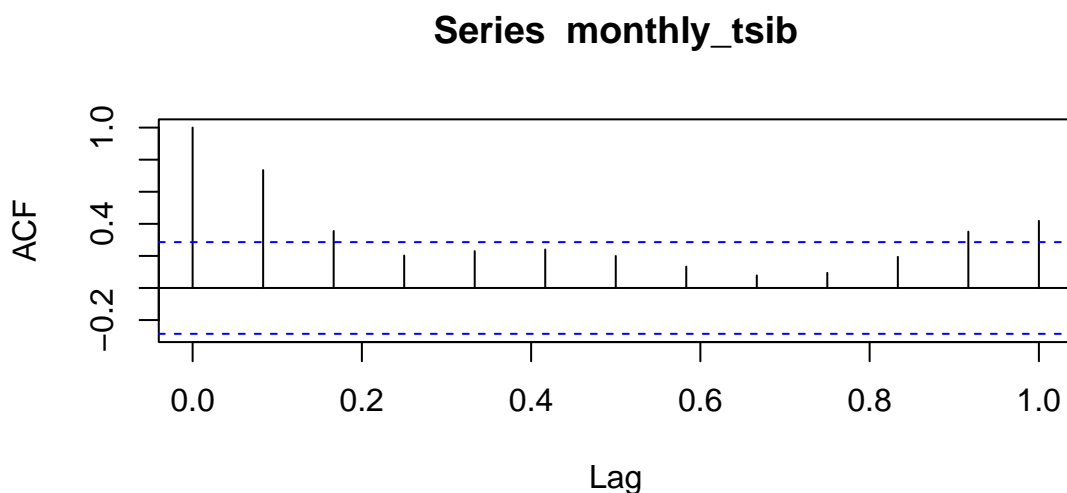
In addition to considering the trend component of the classical decomposition from March 2020 onward, I also complete a hypothesis test to compare the mean monthly participation in the NSLP before and during the Covid-19 pandemic.

Mean monthly participation in the NSLP prior to Covid-19 was 419605973 and mean monthly participation in the NSLP during the Covid-19 pandemic is 122430863.

My null hypothesis is that there is no significant difference between the pre- and during- Covid-19 lunches served, and my alternative hypothesis is that there is a significant different between the conditions. I test the null hypothesis using a Welch's two-sample t test.

Summarize results visually and in words

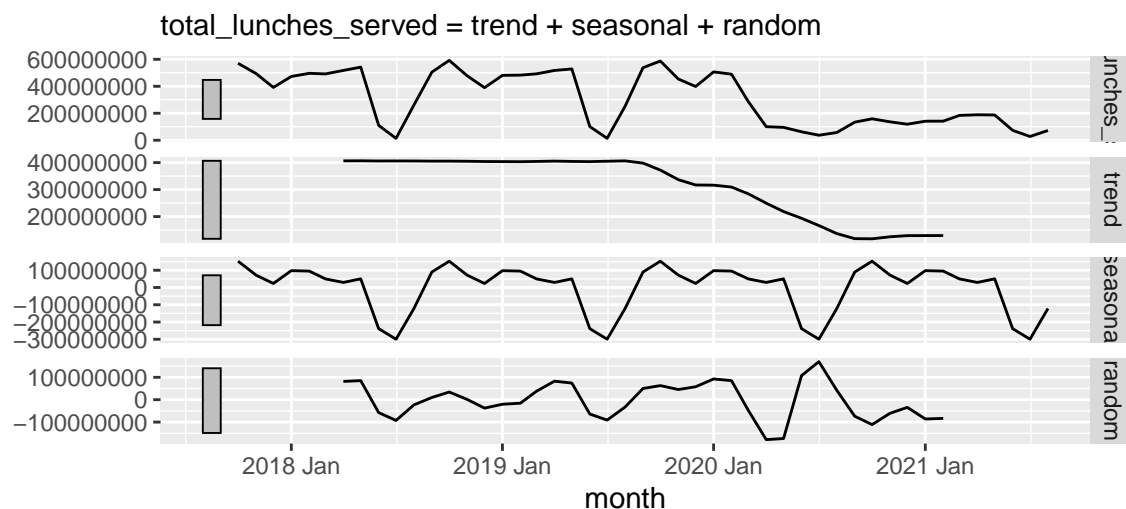
Seasonality of NSLP



The results of the autocorrelation function (ACF) indicate that monthly participation in the NSLP for this month is positively correlated with monthly participation in the NSLP for the next two months, and the prior two months with statistical significance. Participation in all of the other months of the year have a positive correlation with the present month, but not of any statistical significance from correlation of 0, as indicated by the blue line indicating 95% confidence interval.

I wondered if isolating participation data to only students who qualify might change the ACF output, since I would expect students who qualify for school meals and participate rely on NSLP more regularly as an isolated group, but the ACF plot was about the same!

Classical decomposition



The results of my decomposition function autoplot demonstrate a stronger seasonal effect in NSLP participation among all participating students. We can see declines in participation during the summer months, as we would expect but also changes in participation during the school year. I predict that because the data is reported as a monthly total, that there are seasonal decreases in participation during months with school breaks (in winter and spring). Because of this and the domination of the summer impact to seasonality, I am not sure that there is evidence enough to say that the NSLP participating schools and districts feed students differently seasonally during the school year.

Any seasonal differences between school months were drowned out by the seasonal differences between school months and summer months. To further investigate the seasonal differences in meals served, I tried to also

run hypothesis tests comparing total meals served for fall and spring months, fall and winter months, and spring and winter months.

```
## # A tibble: 1 x 10
##   estimate estimate1 estimate2 statistic p.value parameter conf.low conf.high
##   <dbl>     <dbl>     <dbl>     <dbl>   <dbl>     <dbl>     <dbl>     <dbl>
## 1 116728134. 422522363.   3.06e8     1.86  0.0820     15.9  -1.66e7    2.50e8
## # ... with 2 more variables: method <chr>, alternative <chr>

## # A tibble: 1 x 10
##   estimate estimate1 estimate2 statistic p.value parameter conf.low conf.high
##   <dbl>     <dbl>     <dbl>     <dbl>   <dbl>     <dbl>     <dbl>     <dbl>
## 1 70046711. 375840940. 305794229.     1.33  0.197     21.8  -3.91e7    1.79e8
## # ... with 2 more variables: method <chr>, alternative <chr>

## # A tibble: 1 x 10
##   estimate estimate1 estimate2 statistic p.value parameter conf.low conf.high
##   <dbl>     <dbl>     <dbl>     <dbl>   <dbl>     <dbl>     <dbl>     <dbl>
## 1 38540202. 344334430. 305794229.     0.634  0.534     18.1  -8.91e7    1.66e8
## # ... with 2 more variables: method <chr>, alternative <chr>
```

Hypothesis testing of differences in mean meals served by school year season showed no significant differences between school year seasons. I would be curious to run this test if more data were available, as when summer data is excluded, there are only 35 observations in the sample, on the cusp of violating the Central Limit Theorem. With the data I have here, I cannot answer the question “does the the USDA NSLP show seasonality in feeding kids?” for seasons when school is in session.

Impact of Covid on NSLP

The decomposition model used to understand seasonal trends in NSLP participation also indicated a strong downward trend in 2020 and 2021, which seems obviously related to school closures. As a gut check, I conducted a hypothesis test to compare the mean monthly participation in NSLP before and during the pandemic. I was not sure how to define the dates of the pandemic besides the start of nationwide lockdowns in March 2020. I considered it ongoing, since the continuing spread of the Covid-19 virus, lag in access to vaccines in students relative to adults, rules around attendance for students with symptoms or risk of interaction with infected individuals, and decisions by parents to remove students from schools are all ongoing barriers to institutions which facilitate NSLP meals. Therefore, all months including and after March 2020 are classified as “covid” in my data, and otherwise “pre-covid.”

TTest_Component	Results
P Value	0.00000000014
T Score	-8.60000000000
Confidence Interval Lower	-367060915.12000000477
Confidence Interval Upper	-227289304.41999998689

A Welch two-sample t test comparing the means of monthly meals served before and during the pandemic had statistically significant results with a p value of 0. This means there is a mean nonzero difference between the number of kids NSLP served before and during the pandemic and there is a less than 1 percent probability of observing a difference as extreme as this if the null (0 difference) were true. The confidence interval indicates that with 95% confidence we can estimate the range -227289304.42 - -367060915.12 contains the difference in mean participation from before and during the pandemic.

Next steps and future directions

The USDA's data platform only allows me to access NSLP participation data from 2017 onward. As a next step, I might reach out to their data services to try to obtain data from previous years. 47 months of data are *just* enough to satisfy the Central Limit Theorem when we exclude summer data, but I wonder if a larger sample size might be better for illuminating conclusive differences in the mean monthly number of students served between school months by season.

Student access to NSLP is vital for supporting equity in U.S. K-12 schools and helping students to be their best, healthiest selves. Understanding gaps in feeding kids - from regular dips to national emergency program closures - can help policy-makers and on-the-ground stakeholders to make data-informed decisions that reduce identified gaps in access to school meals. Enhancing NSLP's meals served supports millions of hungry students and their families (5). I hope that others continue to make use of the USDA's publicly available national school meal data to continue to improve the ways we understand USDA program operations in schools.

References

- (1) <https://www.ers.usda.gov/topics/food-nutrition-assistance/child-nutrition-programs/national-school-lunch-program/>
- (2) <https://fns-prod.azureedge.net/sites/default/files/resource-files/NSLPFactSheet.pdf>
- (3) <http://bestpractices.nokidhungry.org/sites/default/files/2020-06/Need%20for%20School%20Nutrition%20Budget%20Relief.pdf>
- (4) <https://www.karger.com/Article/Pdf/66399>
- (5) <https://www.ers.usda.gov/webdocs/publications/102076/err-298.pdf?v=3840.6>