Nutrition Intervention Program

By Sarah Little and Lauren Kohls

Table of contents

01
Introduction
NIFA Project

02
Exploratory Analysis

03
Statistical Analysis
T-tests and ANOVAS

04
Multilevel Modeling
Hypotheses/ Findings

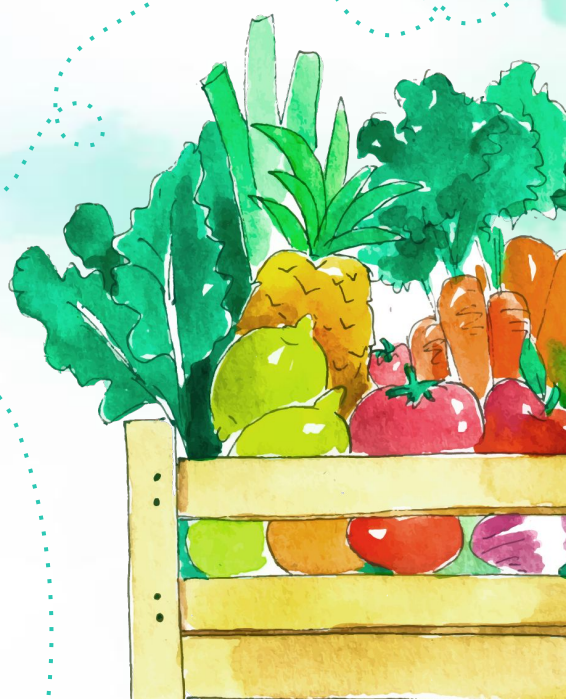
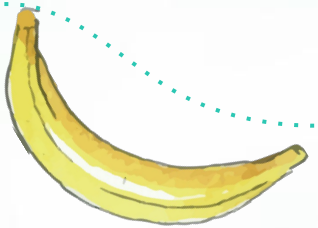
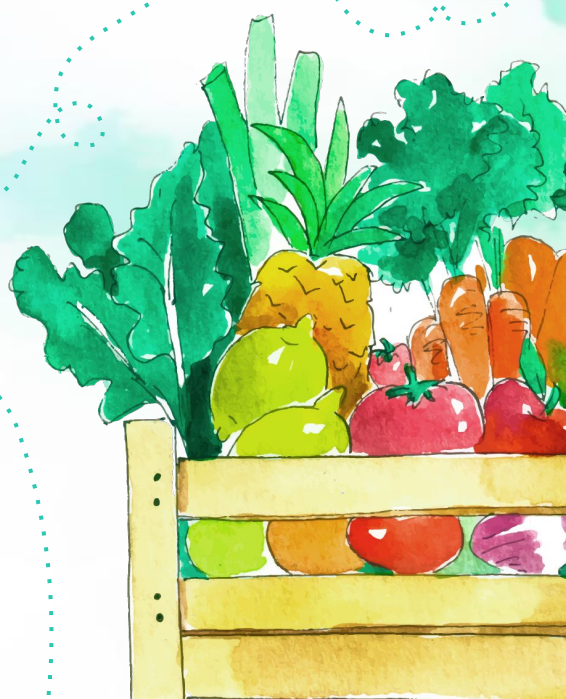


Table of contents

05
Limitations

06
Conclusions





01.

Introduction

The Nutrition Intervention Program

“The Nutrition Intervention Program is a multicomponent nutrition education program designed to empower teachers to improve nutrition literacy and prevent obesity in elementary school students.”

–Research Methods

Levels of Intervention

Professional Development Sessions

Teachers in the Experimental schools have the opportunity to receive lessons on how to deliver nutritional education

Within Classroom Lessons

Students in the Experimental schools receive nutrition focused classes

Study Subjects

Schools

Experimental and Comparison

Teachers

25 teachers selected per school

Students

Approx. 200 Students selected per school

02.
Exploratory Analysis

Multilevel Data

4 schools

School

Teacher ID

Individual Teacher Identification

Individual student identification

Student Code

Level

Type of test depends on level of student

Multilevel Data

NIFA Program Year

Program Year

Teacher Year

Number of Years the Teacher Has Been Involvedt

Individual student Year

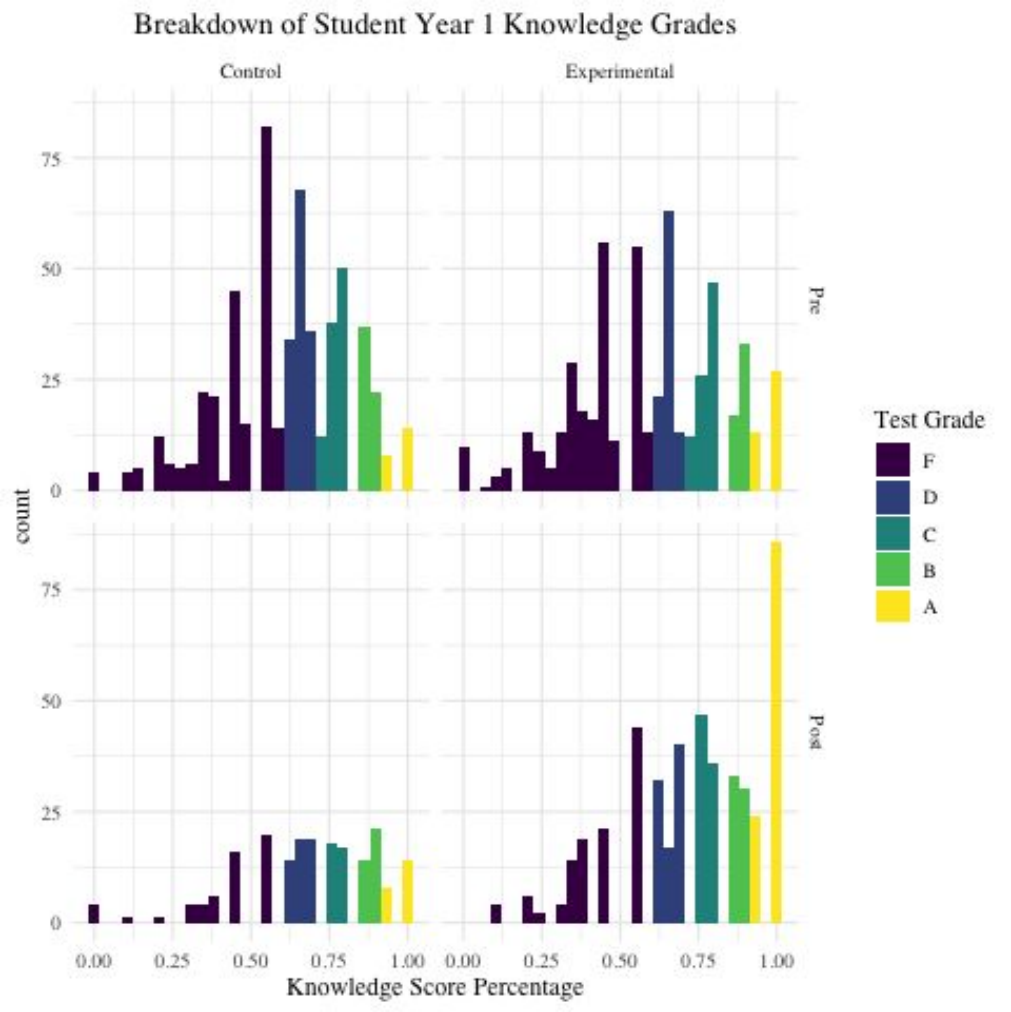
Student Year

Test Exposure

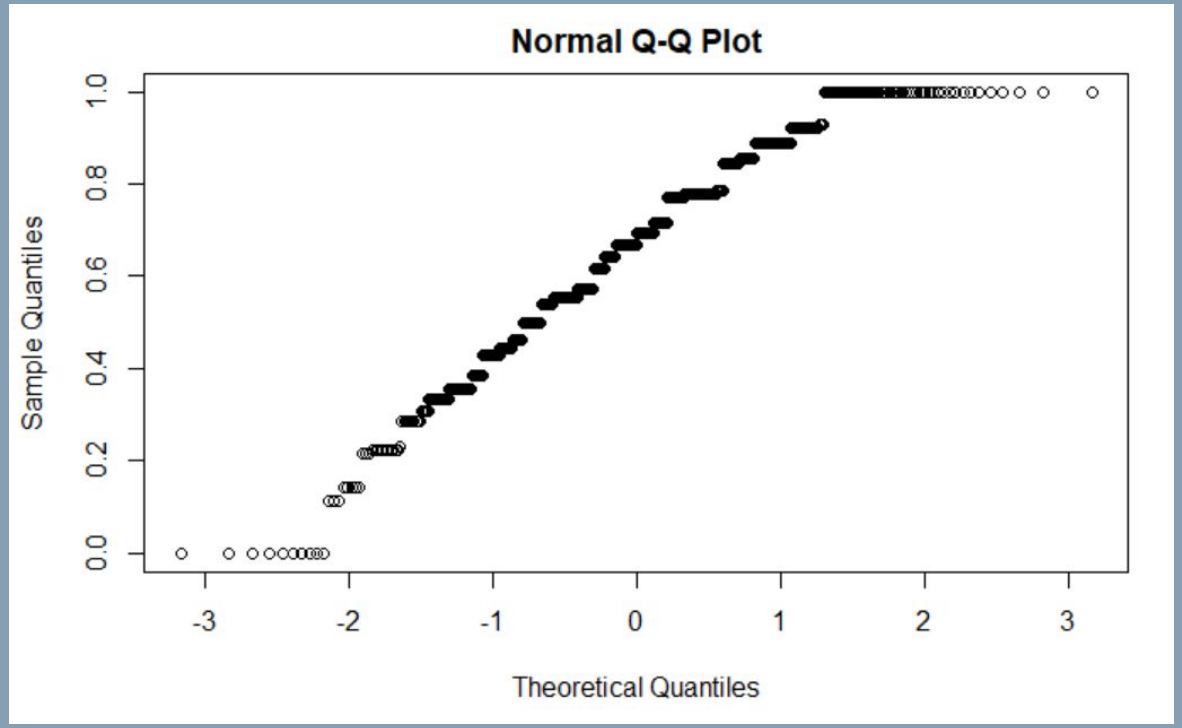
How many times has the student seen the test

Variables of Interest

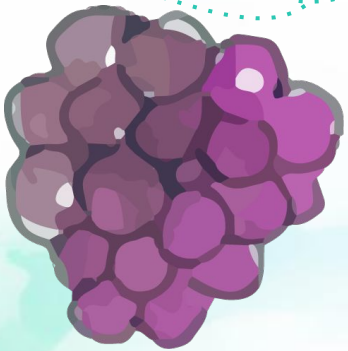
Test Scores, Professional Development Lessons, Lessons Received, Test Exposure, School Type, Levels



- Students are in their first year of the program
- Broken down by Pre and Post test
- Distinct difference in distribution of grades between school types and pre/ post test



- Not Normally Distributed
- 42 Outliers
- Shapiro Wilks Test
- PY1: PRe Knowledge Percent ($n = 325$, $p < .05$)
- PY1 Post Knowledge Percent ($n = 320$, $p < .05$).
- Non-Parametric Tests:
  - Wilcox Ranked Sum
  - Kruskal-Wallis



03.
Statistical Analysis

Teacher Data

- No significant correlation between pre self-efficacy scores and #PD sessions attended among intervention teachers (p=.98, df = 17, t = .025, cor = .006)
- There is a significant correlation between # of PD sessions attended and # of lessons taught among intervention teachers (p = .01, df = 24, cor = 48) *

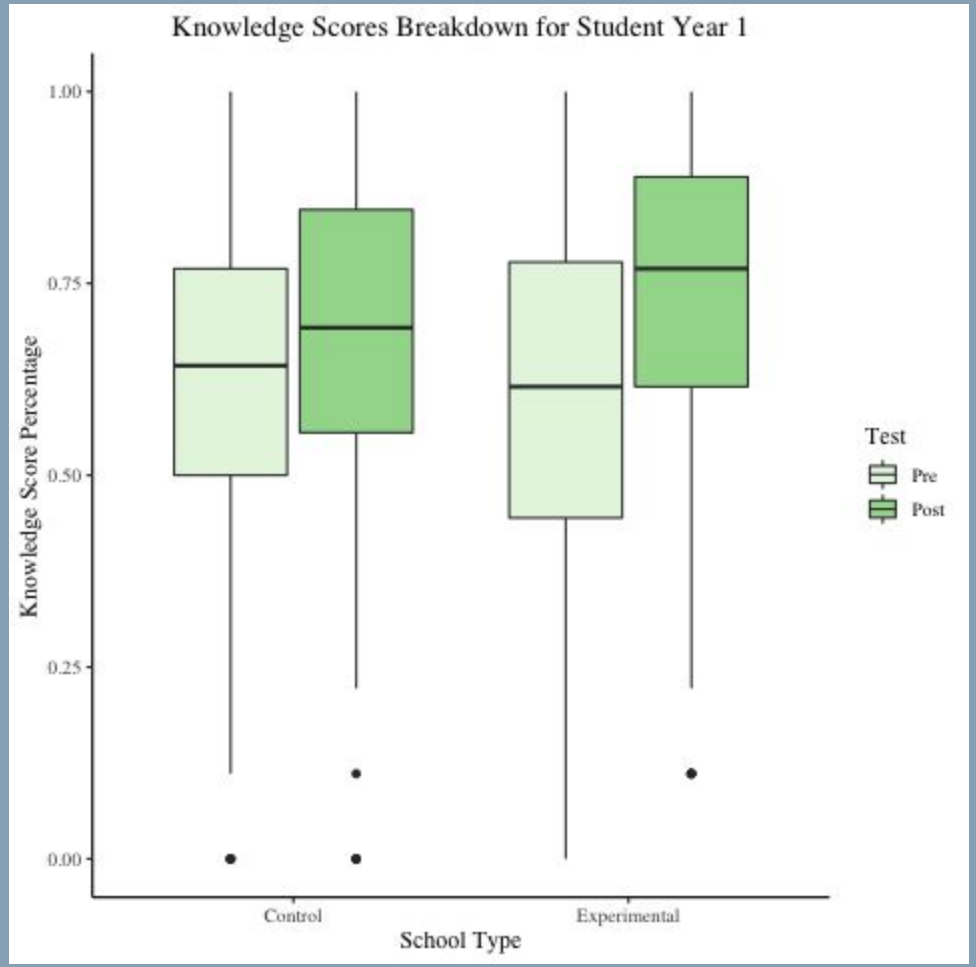
* For this analysis we looked at Experimental School Teachers that attended at least 1 PD Session*

Program Year =1

Student Data

- No significant difference between pre knowledge scores between schools (p=.2, df = 1091) // Student Year 1 Excluding Y3 Data: (p = .5, n = 611)
- There is a significant difference between Post knowledge scores between school type (p = .008 , n= 659)
- There is a significant difference between Knowledge Percent Score Changes between school type (p = 2.08e-5 , n= 659)

Student Year =1

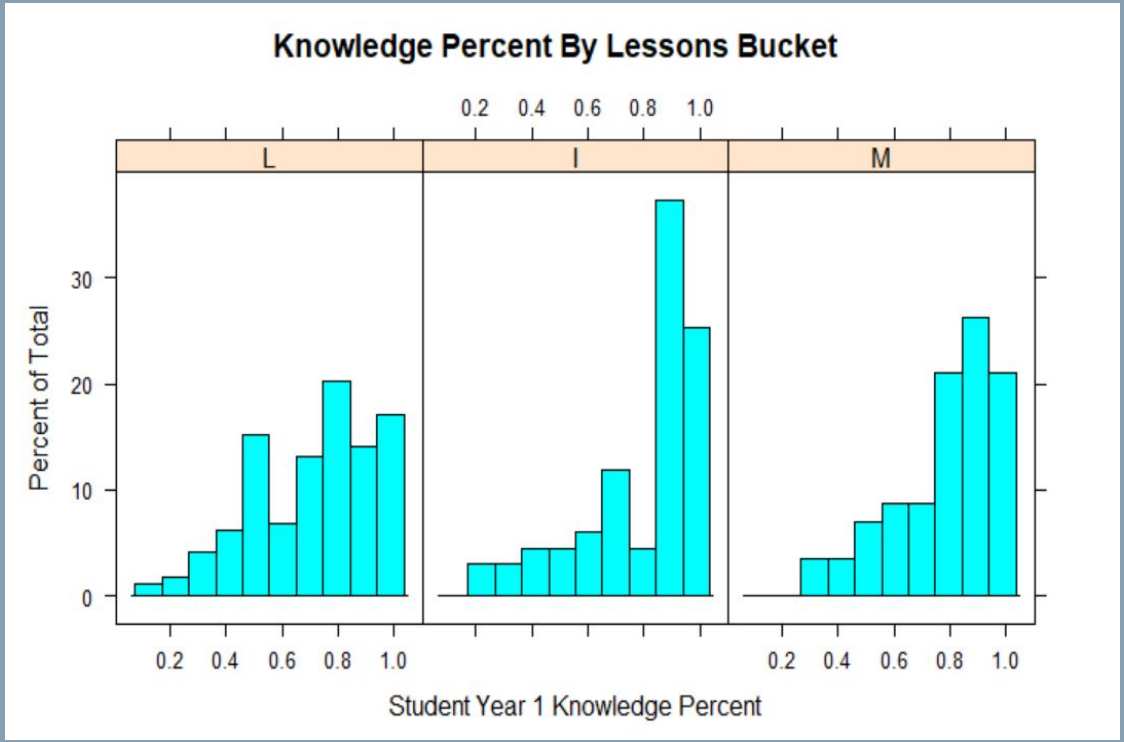


- Students are in their first year of the program
- No significant difference in Student Y1 Pre Knowledge Percents ($p = .16$, $n = 1091$)
- Post Knowledge Percent Score between school types ($p = .008$, $n = 659$)
- Significant differences in Knowledge Percent Score Changes between school types ($p < .01$, $n = 659$)

Change(Multiple Comparison of Means)

- Students who received the intervention (3) had scores that were on average 10% higher than those fewer lessons (0-2)
  - $p < .001$, $df = 2$, $F = 9.66$
- Students that received more than 3 lessons had scores that were 8% higher on average than those who received fewer lessons (0-2)
  - ($p = .01$).
- **There was no statistical** difference in scores for students that received the intervention and those that received more than 3 lessons.

Student Year =1



- Knowledge Percent by lesson bucket
- Statistical Difference between Less and the INtervention
- More and the Less
- No statistical difference between More and Intervention. We therefore recoded our lesson bucket to 0 for less than 3 lessons received
- 1 for more than 3 lessons received

04.
Multilevel Modeling

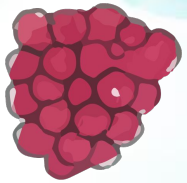
Ordinal Logistic Regression

Ordinal logistic regression (OLR) is used to determine the relationship between a set of predictors and an ordered factor dependent variable.

Ordered	Knowledge Scores are grouped into Grades (A-F)
Random Variables	Code, Teacher ID, Level within School
Fixed Variables	Professional Development, Lessons Received, Gender...

Null Hypothesis

The interventions cause no difference in Knowledge Scores obtained between the Experimental and Control schools.



Alternative Hypothesis

The interventions lead to an increase in the Knowledge Scores obtained by students in the Experimental school.

Imputation

postselfefficacymean	0.53333333
L1CombinedPercent	0.52708333
L1CombinedPercentChange	0.49583333
preselfefficacymean	0.42583333
prewtpersonalhealthsubscore	0.42583333

Starter Models

Null

No fixed effects:

Knowledge Grades ~ 1

Fixed Effects

No random effects:

Knowledge Grades ~ PDSessions + Lessons + PDSessions * Lessons

Random Effects

No fixed effects:

Knowledge Grades ~ 1 + (1|Code) + (1| TeacherID) + (1|Level:School)

ANOVA on Models

anova(Null, Fixed)

```
         no.par    AIC  logLik LR.stat df Pr(>Chisq)
fm.nod        4 7299.3 -3645.7
fm.fixed      6 7228.3 -3608.2  74.999  2  < 2.2e-16 ***
```

anova(Null, Random)

```
         no.par    AIC  logLik LR.stat df Pr(>Chisq)
fm.nod        4 7299.3 -3645.7
fm.rand       7 7034.8 -3510.4  270.48  3  < 2.2e-16 ***
---
```

Model #1

Knowledge Grade ~ PDSessions+ Lessons + Test Exposure + Level + (1|Code) + (1|TeacherID) + (1|Level:Grade)

```
       no.par    AIC  logLik LR.stat df Pr(>Chisq)
fm.nod      4 7299.3 -3645.7
fm.1        9 6962.9 -3472.4  346.45  5  < 2.2e-16 ***
```

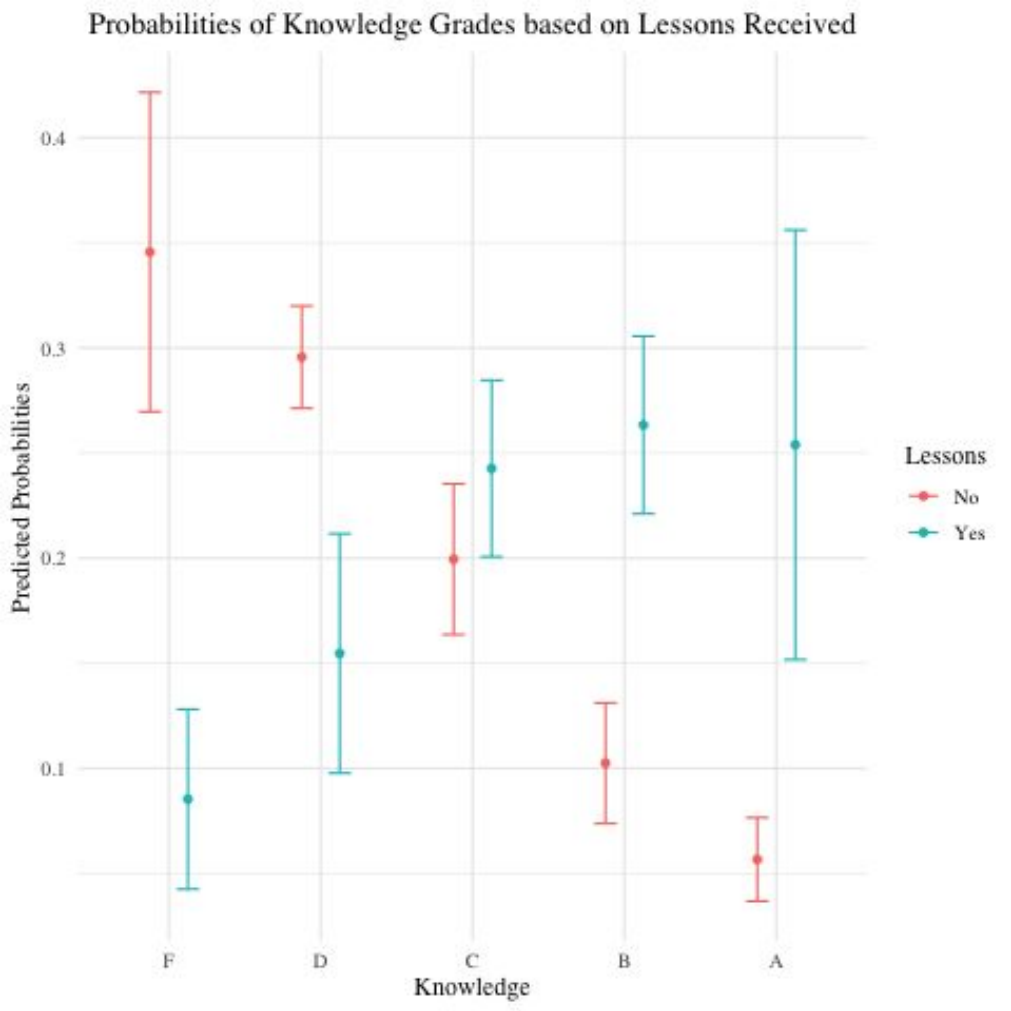
CLMM Assumptions Model 1

1. Dependent variable is ordered.	✔
2. One or more of IV are either continuous, categorical or ordinal.	✔
3. No multicollinearity	Lessons Received and PDSessions violates
4. Proportional odds	PDSessions violates

Log Odds Ratios

KP_brac ~ lessons_1 + (1|Code) +(1|TeacherID)+(1|Level:School))

- β = 1.733 * Lessons Received = 0,1 (change in log odds)
  - Students who received lessons have an estimated coefficient of an increase 1.733 change in log odds in receiving a certain grade.
- α = 2.81 B|A threshold (log odds)
  - The log odds of receiving an A or below, holding all other variables constant is 2.81



- Looking at probabilities instead of log odds
- Estimated probability of a Student receiving an F with < 3 lessons is greater than receiving an F with > 3 lessons
- Estimated probability of receiving an A with > 3 lessons is great than < 3 lessons

Final Model

Knowledge Percent ~ Lessons + Gender + Test exposure * Test + test exposure * Program Year + (1|Code) + (1|TeacherID) + (1|Level:School)

```
       no.par    AIC  logLik LR.stat df Pr(>Chisq)
fm.nod      4 7299.3 -3645.7
fm.6       14 6838.2 -3405.1  481.07 10  < 2.2e-16 ***
```

CLMM Assumptions Final Model

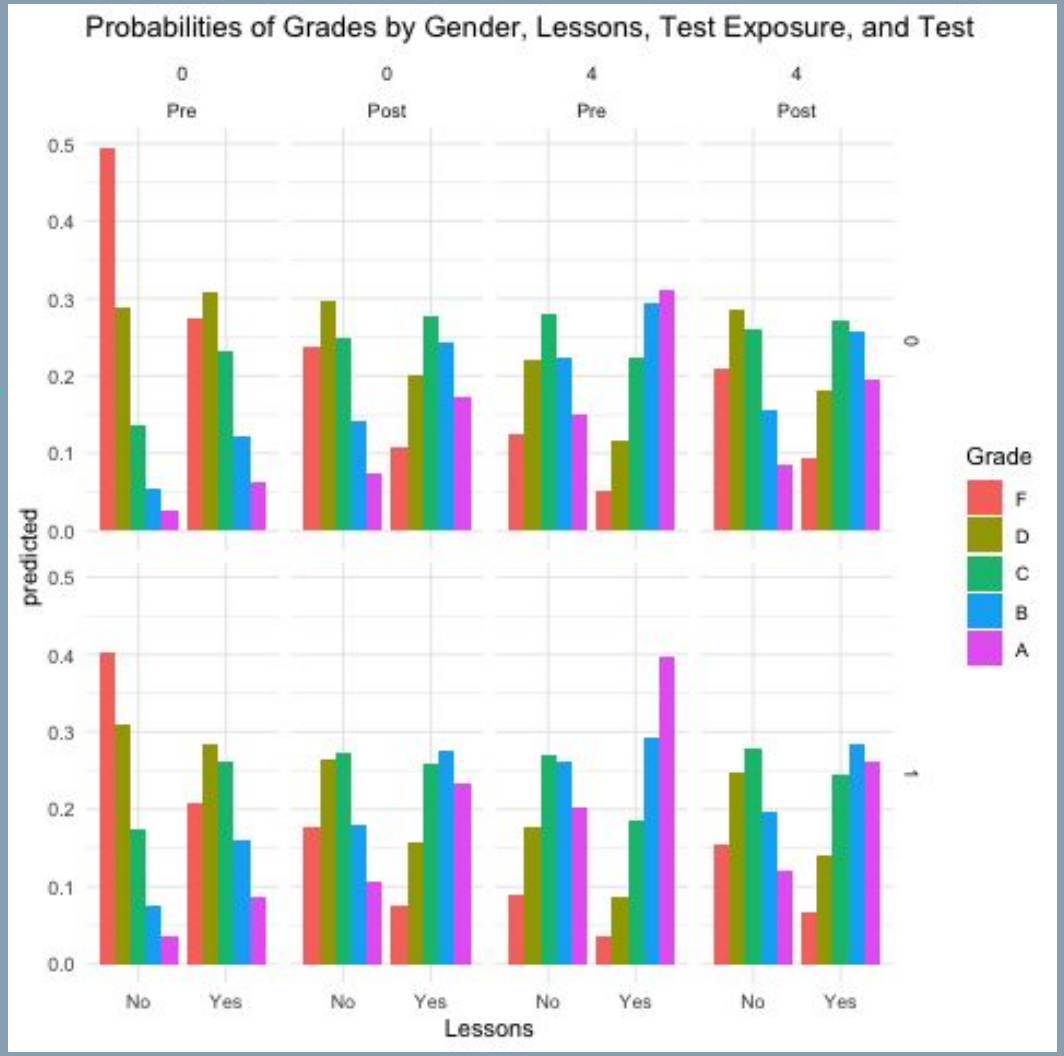
1. Dependent variable is ordered.	✔
2. One or more of IV are either continuous, categorical or ordinal.	✔
3. No multicollinearity	✔
4. Proportional odds	Program_Year violates

Odds Ratios

	exp.coef.fm.6..
F|D	2.1422876
D|C	7.9052421
C|B	25.0324612
B|A	84.9671367
lessons_1Yes	2.5826917
test_exp	3.4672701
Program_Year	1.4785703
TestPost	3.1517943
Gender_1	1.4535507
test_exp:Program_Year	0.6827096
test_exp:TestPost	0.6431342

KP_brac ~ lessons + Gender + test exposure * Test + test exposure * Program Year + (1|Code) + (1|TeacherID) + (1|Level:School)

- The odds of students receiving a grade of **anything but** an A would decrease by 2.58 if the student received lessons.
- The odds of students receiving anything but an A would decrease by 3.46 for every unit increase in test exposure



Example Interpretation:

- Probability of a Student receiving an F with < 3 lessons, test exposure = 0, Gender = Boy, Test = Pre is much higher than the probability of them receiving an A

Findings (odds)

Lessons Received

A change in Intervention from < 3 to >=3 lessons received results in a 2.58 increase in the odds ratio ($p < .01$)

Gender

A change from Boy to Girl results in a 1.45 increase in the odds ratio ($p < .01$)

Test Exposure

One unit increase in test exposure resulting in a 3.47 increase in the odds ratio ($p < .01$)

Test

A change from Pre to Post Test resulted in an 3.15 increase in the odds ratio ($p < .01$).

Model Improvement

- **Comparing *First Model* to *Null Model*** (Nagel Kerkeke Test)
  - *Model #1* returned a pseudo r-squared = .13 compared to the null model
- **Comparing *Final Model* to the *Null Model*** (Nagel Kerkeke Test)
  - *Final Model* returned pseudo r-squared = .20
  - We do see improvement!

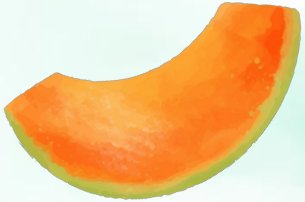
Limitations

Missing Data

Most variables had over 80% missing

COVID-19

Effected the Schools and collection of data



Model Assumptions

Lack of clmm functions

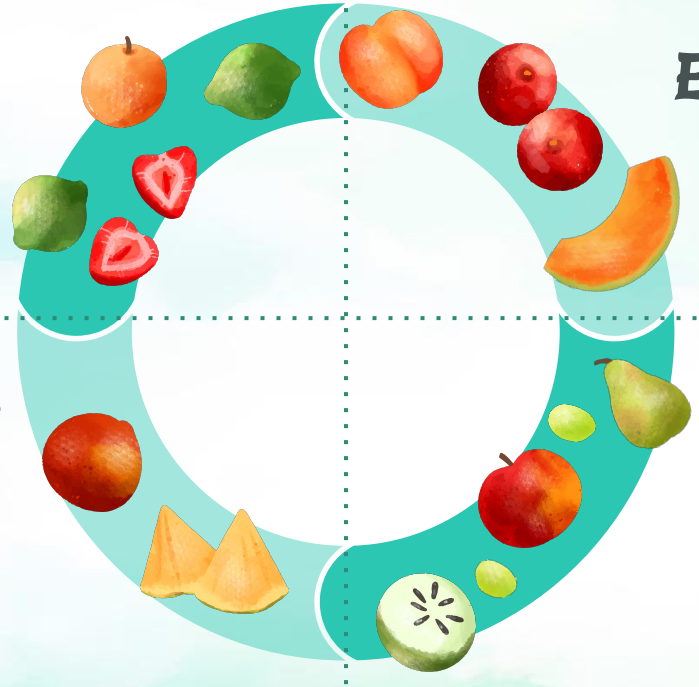
Conclusions

The Data

Multileveled and imputed

Exploratory Analysis

Clear difference between pre and post scores



Statistical Analysis

Lessons Received significant on Knowledge and a correlation between pdsessions and lessons taught

Modeling

Improves when accounting for new variables, random slopes, interactions

End

Bibliographical references

- Agresti, A. (2010). Analysis of ordinal categorical data (2nd ed.), Wiley Series in Probability and Statistics Hoboken, New Jersey: Wiley.
- Christensen RHB (2018). “ordinal—Regression Models for Ordinal Data.” R package version 2018.8-25, URL http://www.cran.r-project.org/package=ordinal/.
- Hoffman, D. L. , & Franke, G. R. (1986). Correspondence analysis: Graphical representation of categorical data in marketing research. Journal of Marketing Research, 23(3), 213–227.
- Kim, T & Park, J. 2019. "More about the basic assumptions of t-tests: normality and sample size. www.ncbi.nlm.nih.gov/pmc/articles/PMC6676026
- Kwak, S Y Kim, J. 2017. "Central limit theorem: the cornerstone of modern statistics. doi: 10.4097/kjae.2017.70.2.144
- Melissa Hawkins, PhD, MHS1; Erin Watts, MPH1; Sarah Irvine Belson PhD2;Anastasia Snelling, PhD, RD1

Thanks

Do you have any questions?

addyouremail@freepik.com
+91 620 421 838
yourcompany.com

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution