Parin Patel

Final Exam Report:

# Purpose:

The purpose of this report is to provide a technical report to the state legislatures office on vaccination data for kindergartens in California in 2013, vaccination rates or five common vaccines provided by the World Health Organization (WHO). The following report will first discuss trends found in the datasets, and then compare the public vs private schools in the state of California with regard to their vaccination information. Finally, the report will conclude with a predicative analysis based on the vaccine data. We hope this information will aid the legislature in shaping future policies with regards to vaccinations of children.

Analysis:

## Part 1- Exploratory Report

Our analysis started by conducting an exploratory review of the vaccination data of private vs public schools. We wanted to see if the school reported vaccinations in 2013, and if there was any difference between private and public schools with regards to reporting vaccinations. You can see in Table1 below that the proportion of public schools that reported Kindergarten vaccinations far outnumbered the number of private schools who did; with 0.76 public schools reporting while only 0.19 private schools reporting. Additionally, the number of public schools that do not report vaccinations is slightly less than the number of private schools who do not; with 0.02 public schools not reporting and 0.03 private schools not reporting . Overall, for both private and public, we do see a larger proportion of schools who report vaccinations than those who do not.

Table 1

|  |  |  |  |
| --- | --- | --- | --- |
| Pub or Private? | Reported? | N | Proportion |
| PRIVATE | N | 252 | 0.03414172 |
| PRIVATE | Y | 1397 | 0.18926975 |
| PUBLIC | N | 148 | 0.02005148 |
| PUBLIC | Y | 5584 | 0.75653705 |

We then reviewed the 5 most common vaccinations reported by the WHO from 1980 to 2017. We wanted to see vaccination rates overtime between the difference vaccines. Figure 1 shows our Time Series plot ,that compares the different vaccination rates. Looking at this, we can see a few patterns. First, it seems that Pol3 (Polio third dose), DTP1, and Hib3 appear to have all stabled off. The Hib3 and DTP1 vaccines started to level off after or around 2010, we would need more data to see if this trend continues. Additionally, the MCV1 vaccine is quite close to having stable rates after 2000, however, because we see a slight dip and return around 2010, we would also need to see data after 2017 for this vaccine to be sure. We are also able to see that the HepB\_BD vaccine has an overall increasing trend overtime, with a peak and then decline after 2010. The Hib3 had an about stable trend until 2010, when it suddenly dipped. Overall, post-2000 MCV1 rates are much more stable than its pre-2000 rates. DTP1 had a variable trend with many spikes and declines, until 2010.

![A close up of a map

Description automatically generated]()

Figure 1

## Part 2- Public vs Private School Comparison

Our next section involved comparing private verse public schools’ vaccinations. See that that overall, between all schools, about 96-100% of private and public schools in California reported vaccinations. We ran a t-test on the public verse private vaccination rates, religious exemption percentage, and medical exemption percentage. The results are showbn below. However, our conclusion is that none of the exemptions had any credible effect.

![A screenshot of a cell phone

Description automatically generated]()

A black sign with white text

Description automatically generated

A screenshot of a cell phone

Description automatically generated

# A screenshot of a cell phone Description automatically generatedA screenshot of a cell phone Description automatically generated

# A screenshot of a tree Description automatically generated

A screenshot of a cell phone

Description automatically generated

![A screenshot of a cell phone

Description automatically generated]()

# Appendix:

library("dplyr")

library('tidyr')

library('ggplot2')

library('scales')

library('ggthemes')

library('ztable')

library('BaylorEdPsych')

install.packages("MCMCpack")

library('MCMCpack')

str(allSchoolsReportStatus)

str(reportSample)

str(usVaccines)

##Part 1: Intrductory/Descriptive Reports

#1. What proportion of public schools reported vaccination data?

#2. What proportion of private schools reported vaccination data?

#Step1:

#adjust the data types for pubpriv variable and the reported variable.

#Reported will be adjusted to ordered factor sice its like a Y or N assignment to the school name.

VPD <- allSchoolsReportStatus %>%

mutate(pubpriv = as.factor(pubpriv) # numeric to factor

, reported = as.ordered(reported)) # numeric to ordered factor

#Step2:

#need to take VPD (public school vaccination data variables) and count the type of school and if reported.

#then will create a proportion table using prop.table. We can also use 'prop=n/sum(n)' as well.

#finall make my proportion table as a data frame (its tbl\_df() properties will be stripped)

VPD\_prop <- VPD %>%

count(pubpriv, reported) %>%

mutate(prop = prop.table(n))

as.data.frame(VPD\_prop)

VPD\_prop

#Outcome:

# pubpriv reported n prop

# 1 PRIVATE N 252 0.03414172

# 2 PRIVATE Y 1397 0.18926975

# 3 PUBLIC N 148 0.02005148

# 4 PUBLIC Y 5584 0.75653705

#Public Schools vs Private that have reported

VPD\_Yes<-subset(VPD\_prop, reported == "Y")

VPD\_Yes

# OUTCOME:

# pubpriv reported n prop

#1 PRIVATE Y 1397 0.189

#2 PUBLIC Y 5584 0.757

#3. Have U.S. vaccinations rates been stable over time?

# ANSWER: No. Exception might be Plo3 after 1995-2000.

#the data has already been convered to a multivariate time series object.

#dates between 1980 and 2017

]

str(usVaccines)

cor(usVaccines)

plot.ts(usVaccines)

airDF <- diff(usVaccines)

plot(airDF)

# 4. Are there any notable patterns in U.S. vaccinations rates over time?

#ANSWER

#DTP1 seems to have sudden spike and declines.

lapply(usVaccines[,1:5], function (x) barplot(table(x)))

var(usVaccines)

#########################################################

#### PART2: Public vs Private School Comparision

#5. Was there any credible difference in overall reporting proportions between public and private schools?

set.seed(12345)

VPD\_prop

#pubpriv reported n

#1 PRIVATE N 252 0.03414172

#2 PRIVATE Y 1397 0.18926975

#3 PUBLIC N 148 0.02005148

#4 PUBLIC Y 5584 0.75653705

#6. Compare overall vaccination rates (allvaccs) between public and private schools. Are there any credible differences?

str(reportSample)

PublicSchools\_vac<-subset(reportSample, pubpriv == "PUBLIC")

PrivateSchools\_vac<-subset(reportSample, pubpriv == "PRIVATE")

mean(PrivateSchools\_vac$allvaccs)

mean(PublicSchools\_vac$allvaccs)

mean(reportSample$allvaccs)

Mean\_plot<-(c(mean(PrivateSchools\_vac$allvaccs),

mean(PublicSchools\_vac$allvaccs),

mean(reportSample$allvaccs)))

hist(reportSample$allvaccs) #frequency of AllVac

###Difference of means for Public vs Private

reportSample$allvaccs[reportSample$pubpriv=='PRIVATE'] #subsetting for pubpriv data private

reportSample$allvaccs[reportSample$pubpriv=='PUBLIC'] #subsetting for pubpriv data public

#t-test to compare means of public and the private

#sample mean calc

PRIVATEMean<-mean( sample(reportSample$allvaccs[reportSample$pubpriv=='PRIVATE'],size=10,replace=TRUE) )

PUBLICMean<-mean( sample(reportSample$allvaccs[reportSample$pubpriv=='PUBLIC'],size=10,replace=TRUE) )

#mean difference

PRIVATEMean - PUBLICMean

#t-test

t.test(reportSample$allvaccs[reportSample$pubpriv=='PRIVATE'] ,reportSample$allvaccs[reportSample$pubpriv=='PUBLIC'])

#Compare the public and the private

install.packages("BEST")

library("BEST")

pubprivBEST <- BESTmcmc(reportSample$allvaccs[reportSample$pubpriv=='PRIVATE'] ,reportSample$allvaccs[reportSample$pubpriv=='PUBLIC'])

#plot

plot(pubprivBEST, main='AllVaccs, Difference of Means: Public vs Private ', )

#7. Compare medical exemptions between public and private schools. Are there any credible differences?

###Difference of means for Public vs Private

reportSample$medical[reportSample$pubpriv=='PRIVATE'] #subsetting for pubpriv data private

reportSample$medical[reportSample$pubpriv=='PUBLIC'] #subsetting for pubpriv data public

#t-test to compare means of public and the private

#sample mean calc

PRIVATEMedicalMean<-mean( sample(reportSample$medical[reportSample$pubpriv=='PRIVATE'],size=10,replace=TRUE) )

PUBLICMedicalMean<-mean( sample(reportSample$medical[reportSample$pubpriv=='PUBLIC'],size=10,replace=TRUE) )

#mean difference

PRIVATEMedicalMean - PUBLICMedicalMean

#t-test

t.test(reportSample$medical[reportSample$pubpriv=='PRIVATE'] ,reportSample$medical[reportSample$pubpriv=='PUBLIC'])

#Compare the public and the private

pubprivMedicalBEST <- BESTmcmc(reportSample$medical[reportSample$pubpriv=='PRIVATE'] ,reportSample$medical[reportSample$pubpriv=='PUBLIC'])

#plot

plot(pubprivMedicalBEST, main='Medical,Difference of Means-Public vs Private', )

#8. Compare religious/belief exemptions between public and private schools. Are there any credible differences?

###Difference of means for Public vs Private

reportSample$religious[reportSample$pubpriv=='PRIVATE'] #subsetting for pubpriv data private

reportSample$religious[reportSample$pubpriv=='PUBLIC'] #subsetting for pubpriv data public

#t-test to compare means of public and the private

#sample mean calc

PRIVATEReligiousMean<-mean( sample(reportSample$religious[reportSample$pubpriv=='PRIVATE'],size=10,replace=TRUE) )

PUBLICMReligiousMean<-mean( sample(reportSample$religious[reportSample$pubpriv=='PUBLIC'],size=10,replace=TRUE) )

#mean difference

PRIVATEReligiousMean - PUBLICMReligiousMean

#t-test

t.test(reportSample$religious[reportSample$pubpriv=='PRIVATE'] ,reportSample$religious[reportSample$pubpriv=='PUBLIC'])

#Compare the public and the private

pubprivReligBEST <- BESTmcmc(reportSample$religious[reportSample$pubpriv=='PRIVATE'] ,reportSample$religious[reportSample$pubpriv=='PUBLIC'])

#plot

plot(pubprivReligBEST, main='Religious,Difference of Means-Public vs Private', )