Coronavirus in the Southeast

STAT 3480 Consulting Project

Ishaan Dey & Tal Dunne

11/23/2020

Executive Summary

Introduction

Project Description

- Restate the client's research problem (for clarification)
- Put the data and the research question in context.
- Why is the client's research relevant?
- Has this research question been studied previously?

Data Summary & Discussion

- How was the data collected? – What is the sampling/ experimental design? – What are some potential biases?
- What variables are measured? What are the units?
- How large is the dataset?
- Some basic summary statistics and graphs of the data should be presented. (These basic distributions will be relevant for the methodology.)

Methodology

- What variables are measured? What are the units?
- How large is the dataset?
- Some basic summary statistics and graphs of the data should be presented. (These basic distributions will be relevant for the methodology.)

Results

- Test statistics
- P-values
- Confidence intervals
- Inference decisions (conclusions

Discussion

- Interpret and discuss the significance of the analysis results.
- Substantiate any conclusions with results.
- Tell the story of how you answered your client's research question(s).
- Discuss the limitations of the study (which can pertain to data, data collection, or methodology).
- Make sure to be delicate – the client doesn't want to hear what a bad job they've done.

Conclusion

- The goal of this section is to address the client's bottom line: what does your progress mean for his/her research problem?

Appendix

Code

```r
library(readr)
library(dplyr)
```

```r
covid_data <- readr::read_csv("covid-data.csv") %>% as_tibble()
covid_data$Mask.Mandate <- dplyr::if_else(covid_data$Mask.Mandate.Start.Date == 'None',
                                          true=0, false=1) %>% as.factor()
covid_data %>% head()
```

```
## # A tibble: 6 x 8
##   State Total.Cases Total.Deaths Death.Rate.Pct Population.Esti~ Cases.Per.1000
##   <chr>       <dbl>        <dbl>          <dbl>            <dbl>          <dbl>
## 1 Alab~      186437         2892           1.55          4903185           38.0
## 2 Arka~      107679         1857           1.72          3017804           35.7
## 3 Flor~      776215        16505           2.13         21477737           36.1
## 4 Geor~      353372         7844           2.22         10617423           33.3
## 5 Kent~       99637         1428           1.43          4467673           22.3
## 6 Loui~      184724         5872           3.18          4648794           39.7
## # ... with 2 more variables: Mask.Mandate.Start.Date <chr>, Mask.Mandate <fct>
```

Big Q1

```r
mask.sample1 <- covid_data %>%
  filter(Mask.Mandate == 1) %>% pull(Cases.Per.1000)
mask.sample2 <- covid_data %>%
  filter(Mask.Mandate == 0) %>% pull(Cases.Per.1000)
```

```r
wilcox.test(mask.sample1, mask.sample2, alternative="less")
```

```
##
##  Wilcoxon rank sum exact test
##
## data:  mask.sample1 and mask.sample2
## W = 14, p-value = 0.183
## alternative hypothesis: true location shift is less than 0
```
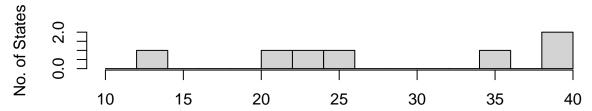
Big Q2

```r
case.rate <- covid_data$Cases.Per.1000
death.rate <- covid_data$Death.Rate.Pct
```

```r
r.obs<-cor(rank(case.rate), rank(death.rate))

source("http://www4.stat.ncsu.edu/~lu/ST505/Rcode/functions-Ch5.R")
perm.r<-perm.approx.r(rank(case.rate),rank(death.rate),2000)

p.upper<-mean(perm.r >= r.obs)
p.lower<-mean(perm.r <= r.obs)
p.two.sided<-mean(abs(perm.r) >= abs(r.obs))

paste("Spearman's Correlation:", round(r.obs,4)) %>% print()
```

```
## [1] "Spearman's Correlation: 0.3022"
```

```r
paste("p-value:", round(p.upper,3)) %>% print()
```

```
## [1] "p-value: 0.154"
```

```r
tau.obs<-getTau(case.rate,death.rate)

perm.tau<-perm.approx.tau(case.rate,death.rate,2000)

p.upper.tau<-mean(perm.tau >= tau.obs)
p.lower.tau<-mean(perm.tau <= tau.obs)
p.two.sided.tau<-mean(abs(perm.tau) >= abs(tau.obs))

paste("Kendall's Tau:", round(tau.obs,4)) %>% print()
```

```
## [1] "Kendall's Tau: 0.2564"
```

```r
paste("p-value:", round(p.upper.tau,3)) %>% print()
```

```
## [1] "p-value: 0.124"
```

Figures

Figure 1. Distribution of Case Rates (per 1000), by State Mask Mandates

Mask Mandated

No. of States

Case Rate Per 1000

Mask Not Mandated

No. of States

Case Rate Per 1000