CSUEB Primary Care and Counseling Services 2022 Statistical Analysis

Author: Dan Hoang, Rahman Baluch

11-07-2022

Contents

I. Purpose of the Study

The purpose of this paper is to strengthen the current study of Dr. Spencer about the utilization rate of our CSUEB health center and the impact of COVID by using time-series analysis and statistical test.

According to Dr. Spencer, number of appointments in all outpatient clinics, including student health centers, have dropped across the country secondary to COVID. Even prior to COVID, studies show that the majority of students have an interest in health, but there is underwhelming use of Student Health Services across the country. An intervention from CSUEB staff has been implemented in attempts to increase the utilization rate of SHCS by reaching out to students via the student health portal reminding them of the medical services available on campus. This paper will go through exploratory data analysis and ANOVA test to analyze the effect of COVID and the effect of interventions upon the CSBEB Primary Care and Counseling Services.

II. Data Description.

Data was collected from Medicat between March 2018 to May 2019 and from PNC between June 2019 to June 2022 prior to data analysis and writing of this report. Dataset has dimension 178 rows and 6 columns (variables). Below are descriptions of each variable.

- `date`: time visits healthcare services.
- `departmentName`: name of departments including counselling, primary care, laboratory and nursing.
- `checkins`: total visits healthcare services.
- `numvisits_perprovider`: total visits healthcare services per provider.
- `pandemic`: pre pandemic and post pandemic.
- `intervention`: several months have been implemented intervention.

III. Exploratory Data Analysis.

There is an unbalanced distribution of number visits through all the months of the year. Most students use health services in middle of Spring and middle of Fall and less on Summer and Winter break, which is normal since students mostly come to campus on two main semesters.

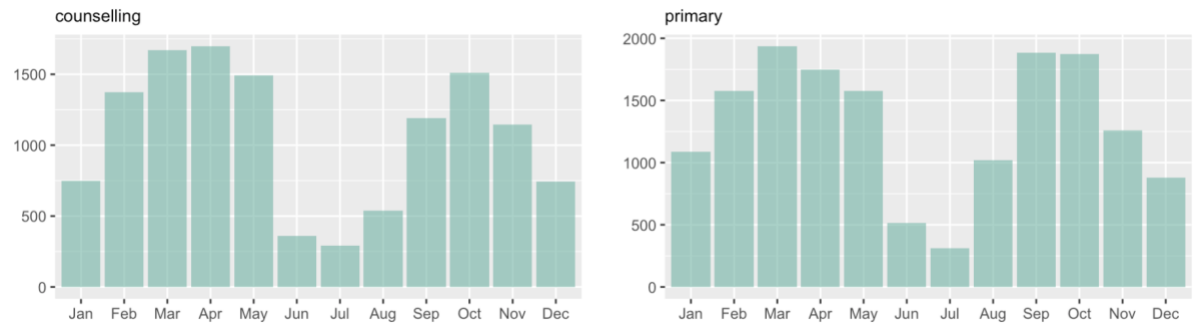

Figure 1. Visits distribution through months.

For days in week, there is not much variation on distribution, except for Wednesday which is has the lowest number visit in both counselling and primary care. Tuesday and Thursday are usually busy than the other days.

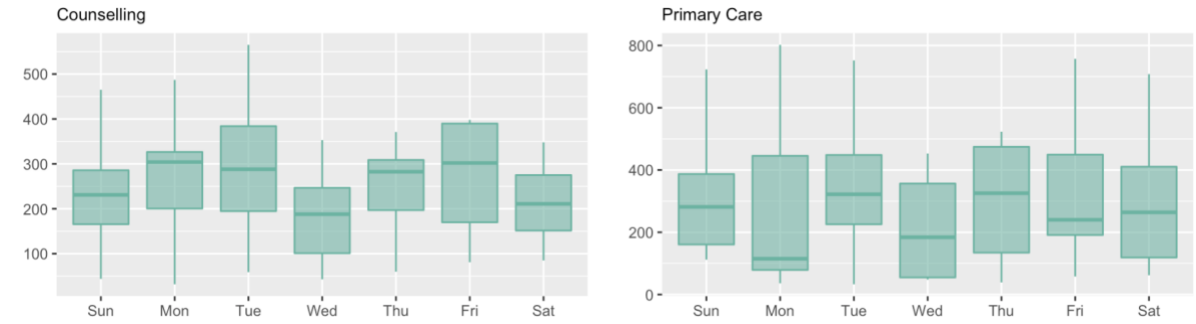

Figure 2. Visits distribution through week days.

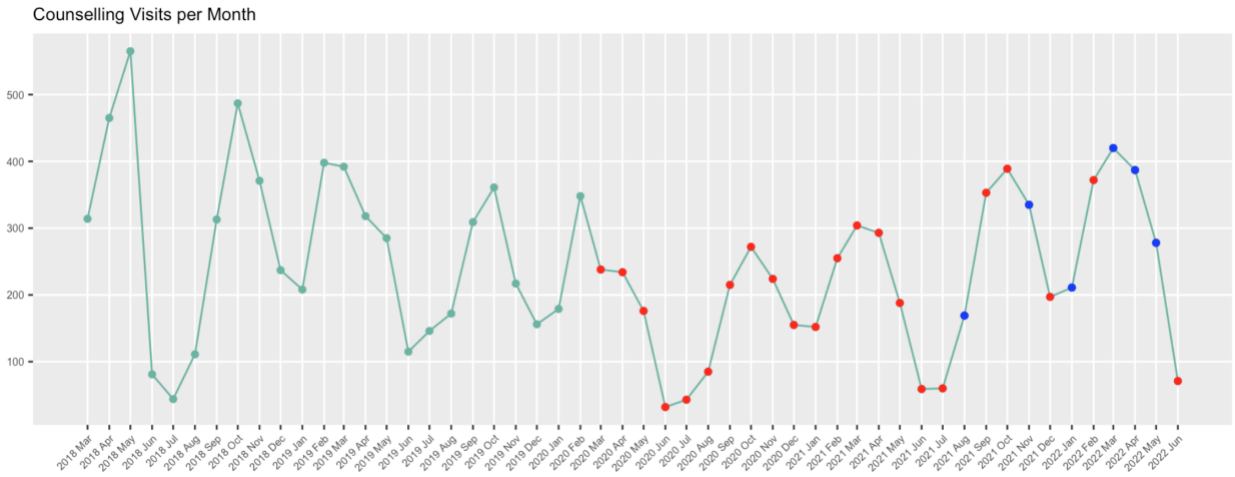

Figure 3. Counselling Visits per Month (red: Covid, blue: Intervention).

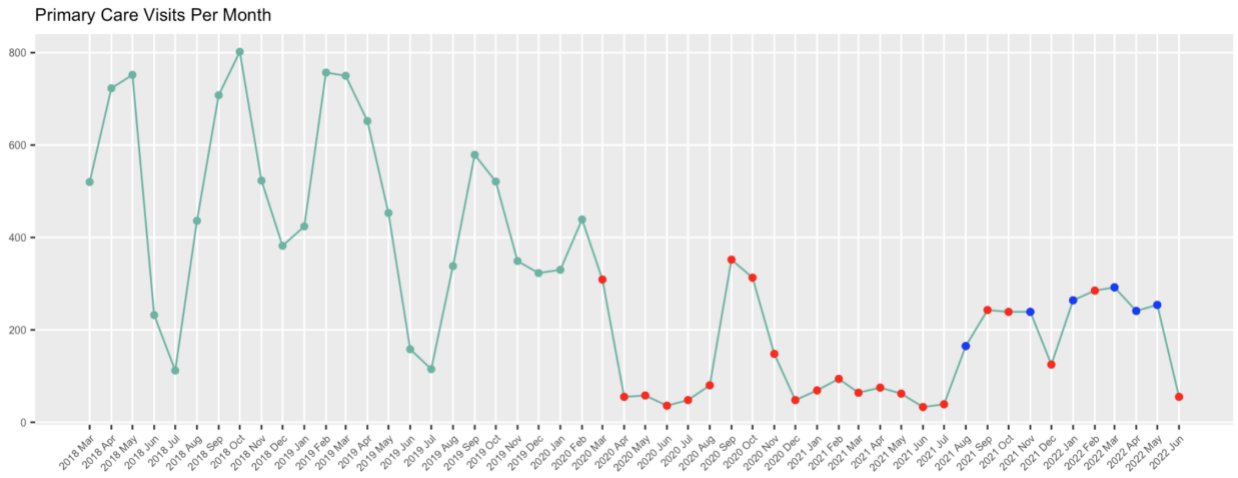

Figure 4. Primary Care Visits per Month (red: Covid, blue: Intervention).

From time-series plot, red points are the Pandemic period, blue points are the months has been implemented intervention. There is not much difference between number visits Counselling pre-pandemic and post-pandemic, however, in Primary Care, there is a visually significant down trend on number visits after Covid. For intervention, it is not enough to make a conclusion since the trend is not clear in the plot, a statistical test would be made to confirm the effect.

IV. ANOVA Analysis.

To determine whether the pandemic Covid or the Intervention has any significant effect, an ANOVA model has been tested. The ANOVA model provides an indication if the means visits count are statistically different between two groups.

a. Pandemic Effect.

Formally, the hypothesizes are as follows:

- H0 (null): There is no significant difference on mean number visits of healthcare services over pre and post-pandemic period.
- H1: There is significant difference on mean number visits of healthcare services over pre and post-pandemic period.

```
#Counselling
anova(lm(Checkins ~ pandemic, data = counsel1))
```

```
## Analysis of Variance Table
##
## Response: Checkins
##           Df Sum Sq Mean Sq F value Pr(>F)
## pandemic   1  38267   38267  2.4858 0.1212
## Residuals 50 769733   15395
```

```
#Primary Care
anova(lm(Checkins ~ pandemic, data = primary1))
```

```
## Analysis of Variance Table
##
## Response: Checkins
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## pandemic   1 1332002 1332002  50.177 4.442e-09 ***
## Residuals 50 1327315   26546
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5. ANOVA result Pandemic Effect

The statistic result show that Primary Care has p-value = 0 < 0.05, which means with 95% confidence level, we can conclude there is significant effect of pandemic Covid on number visits on Primary Care services. In contrast, for Counselling, it fails to reject the null hypothesis, this can be interpreted that there is no significant effect of pandemic Covid on number visits on Counselling services, this result can be seen clearly from Figure 3.

b. Intervention Effect.

Unlike the pandemic effect, there is a huge unbalance between the number of months that implemented intervention and the number of months that do not have intervention (6 versus 46). Authors have consulted Dr. Li Zou about unbalanced data and the proper approach to get a strong ANOVA inference. “Since our data is observational data, we want to use matching to control potential confounders. In order to control the factor of time (month), we should only select the non-intervention data from the same months instead of using all month data” (Dr. Li Zou, 2022). From this foundation, non-intervention group are chosen from the same months that interventions have been implemented. Another problem that will affect the analysis is that if we use year the 2021 data to compare 2022 data, the pandemic effect may overlap with the intervention effect, to make it balanced, the year 2020 will be added to the test, in the graph below, non-intervention group is yellow points, intervention group are blue points. This is not the ideal condition to conduct ANOVA however we can observe many scenarios.

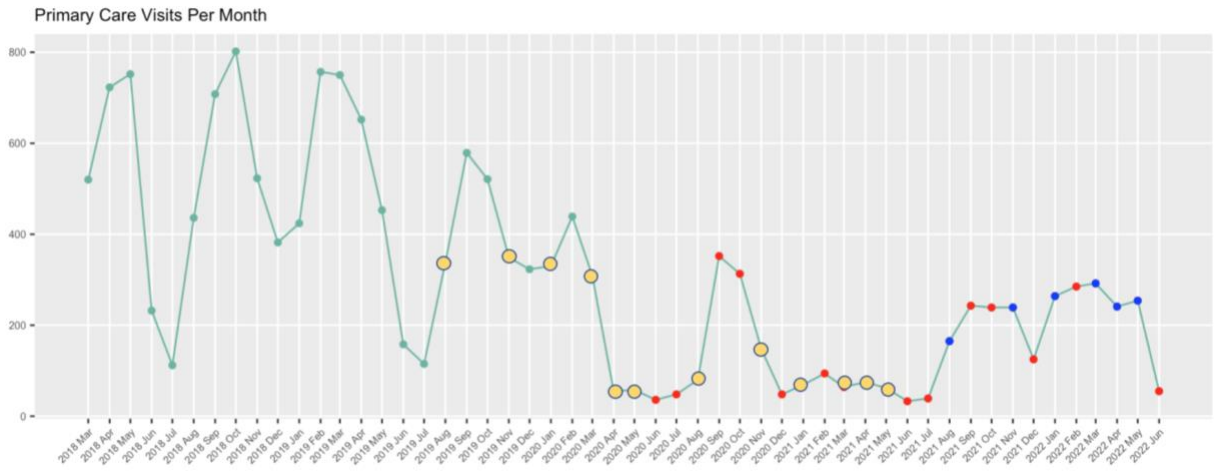

Figure 6. Months collected for ANOVA (yellow: non-intervention, blue: intervention).

The hypothesizes are as follows:

- H0 (null): There is no significant difference on mean number visits of healthcare services with and without intervention period.
- H1: There is significant difference on mean number visits of healthcare services over with and without intervention period.

```
#Counselling
Anova(lm(Checkins ~ intervention, data = counsel_in), type = 2)
```

```
## Anova Table (Type II tests)
##
## Response: Checkins
##              Sum Sq Df F value Pr(>F)
## intervention  20314  1   1.751  0.207
## Residuals    162418 14
```

```
#Primary Care
Anova(lm(Checkins ~ intervention, data = primary_in), type = 2)
```

```
## Anova Table (Type II tests)
##
## Response: Checkins
##              Sum Sq Df F value Pr(>F)
## intervention  26298  1  2.2164  0.156
## Residuals    189846 16
```

Figure 6. ANOVA result Intervention Effect

The statistic result show that s both Primary Care and Counselling we fail to reject the null hypothesis, which means with 95% confidence level, we can conclude there is no significant effect of Intervention from CSUEB staff on number visits on Primary Care and Counselling services.

V. Conclusion

After the data analysis, the obtained results demonstrate some insight of the CSUEB healthcare center visits and some trends are interesting to find out, which is helpful for improving utilization on SHCS. The effect of pandemic Covid on CSUEB on Primary Care services has been confirmed to be significant after testing ANOVA.

Due to the limit of the observational data as well as possible overlapped effect from pandemic period, it appeared to be not an ideal condition to conduct a strong statistical inference for intervention effect, and base on the current data, there is not enough statistical evidence to confirm the effect from intervention on CSBEB Primary Care services. Authors suggest to implement more interventions to detect from pandemic effect and acquire sufficient inputs for further analysis.

References

1. Spencer Wong & Shauna Hong, (2022), Student Health and Counseling Services (SHCS) Utilization Study 2022.
2. Li Zou (2022).

Code Appendix

For supplementary R script, visit: https://github.com/atndan/-SHCS-Utilization-Study-CSUEB-2022