Museums in Singapore: Visitor Insights

Clarence TAY Cheng Wei

Heranshan S/O Subramaniam

LIU Chentao

Melissa TAN May Hua

Ranice TAN Hui Qi

Agenda

Introduction

Objective

Data

R Shiny Demo

Analysis

Summary

Future Work

Introduction

"Our national museums and heritage institutions are cornerstones of the preservation and celebration of our heritage... This is why it is vital for our museums and institutions to continue being **accessible** to all, for all. Our shared goal is to make the Singapore museum experience **enjoyable** and **inclusive**... Ultimately, this ensures that our heritage remains a source of reflection and inspiration for us all."

-National Heritage Board (NHB), Singapore

Objective

- Conduct exploratory data analysis on NHB museum visitors over the past 4 years
- Examine the extent to which museum visits have been

  (i) accessible

  (ii) enjoyable

  (iii) inclusive
- Evaluate the impact of COVID-19 on museum visitorship.

Data Sources

- Provided by the NHB
- Comprises of data from 2017 to 2020
  - Includes 6 national institutions
- Excludes tourists and foreign visitors.
- Includes visitor demographics info and details of their visits

Data field	Description	How it will be used
MDAS ID	Unique identifier for each visitor	To identify repeat visitors
Date	Date of visit (dd/mm/YY)	To project peak/non-peak periods within a day, and throughout the year
Time	Time of visit	To project peak/non-peak periods within a day
Museum	Name of museum	To examine different visitorship and demographic patterns for each museum
Age	Age of visitor	For demographic analysis
Gender	Gender of visitor	For demographic analysis
Race	Race of visitor	For demographic analysis
Marital Status	Marital status of visitor	For demographic analysis
Region, Planning Area, Subzone	Residential location of visitor	To examine distribution of visitors' residential location (origin)

National Museum of Singapore **(NMS)**

Asian Civilisations Museum **(ACM)**

The Peranakan Museum **(TPM)**

Malay Heritage Centre **(MHC)**

The Indian Heritage Centre **(IHC)**

Sun Yat Sen Nanyang Memorial Hall **(SYS)**

Data Preparation

- Missing Data
  - Rows with missing demographics data and visit time were dropped
- Additional derived variables to facilitate analysis
  - Repeat visitors were derived from MDAS ID unique identifier field
  - Visitors' age was calculated from their year of birth and visit date
  - Museums' location were derived based on latitude and longitude data
  - Visitor Date, Time fields were transformed into datetime objects in R using lubridate library
- Additional dataset:
  - National population statistics from data.gov.sg were used to analyse the demographic patterns of museum visitors against national averages

Singapore Museums

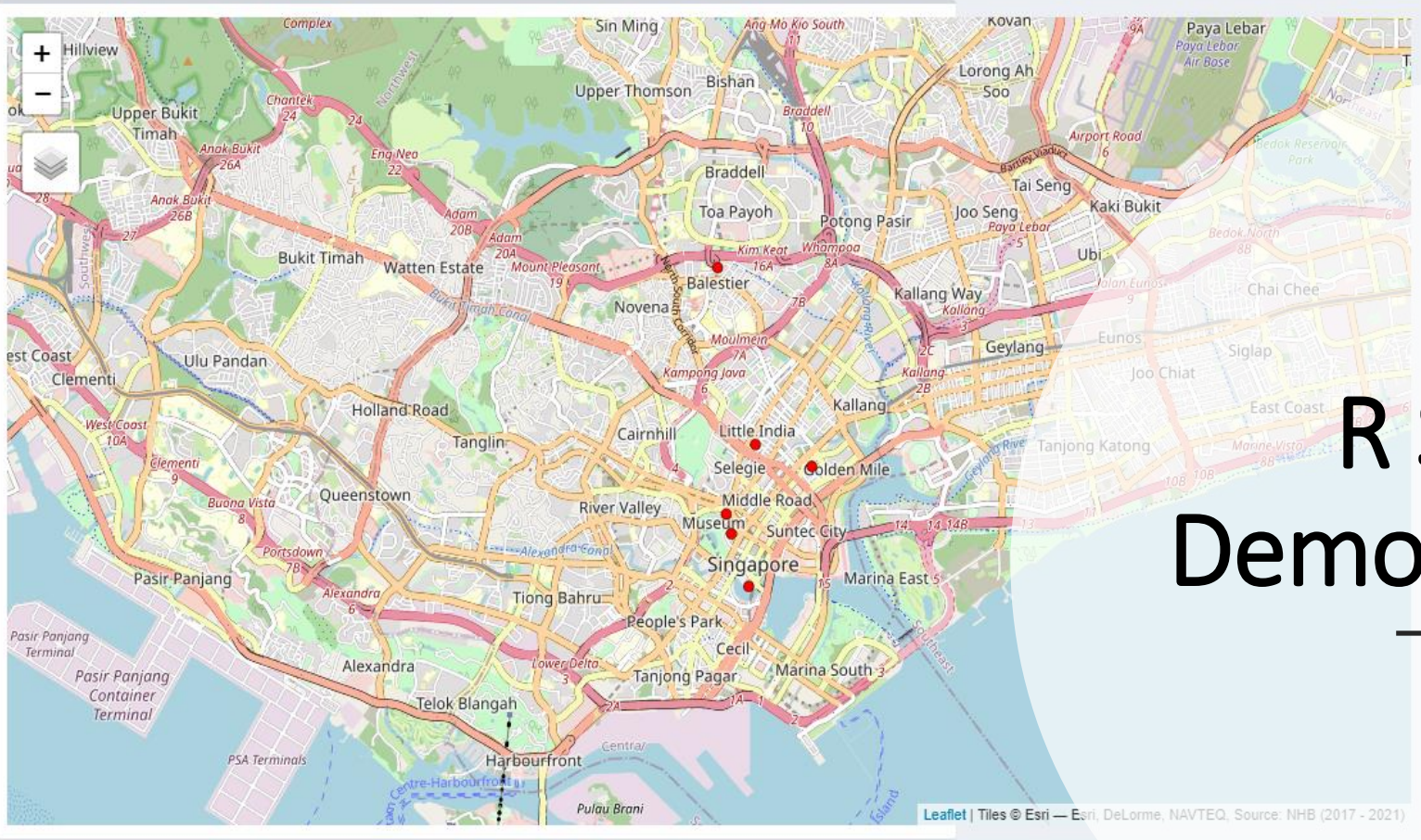
- Home
- General Visitorship Statistics
- Visitor Profile
- Peak Periods
- Origin

Museums in Singapore: Visitor Insights

"Our national museums and heritage institutions are cornerstones of the preservation and celebration of our heritage...
This is why it is vital for our museums and institutions to continue being accessible to all, for all.
- National Heritage Board (NHB), Singapore

Use this app to find out more about the visitor patterns and profile at each museum.

Should you wish to visit, check out the peak timings to avoid. Enjoy your visit.

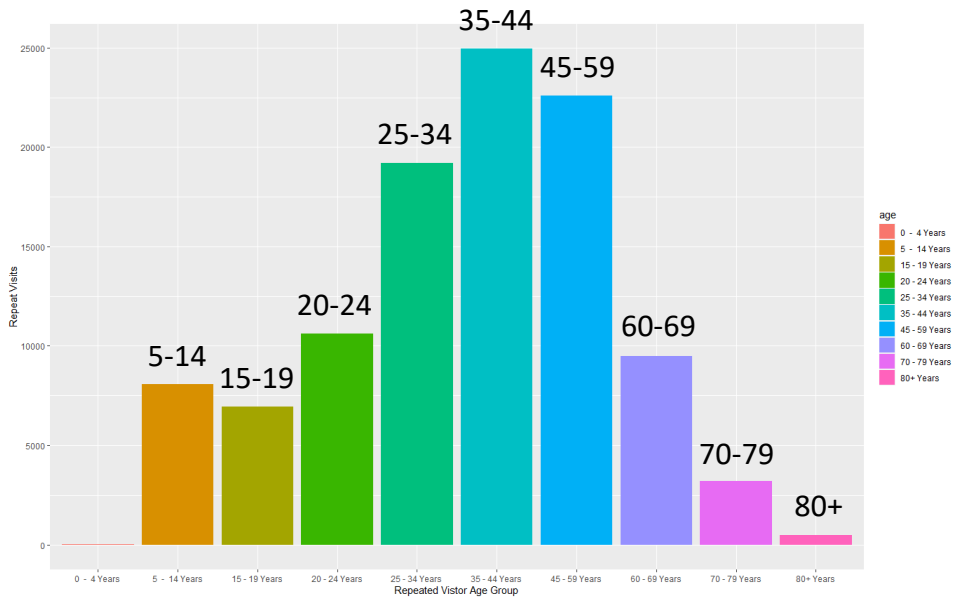


R Shiny Demonstration

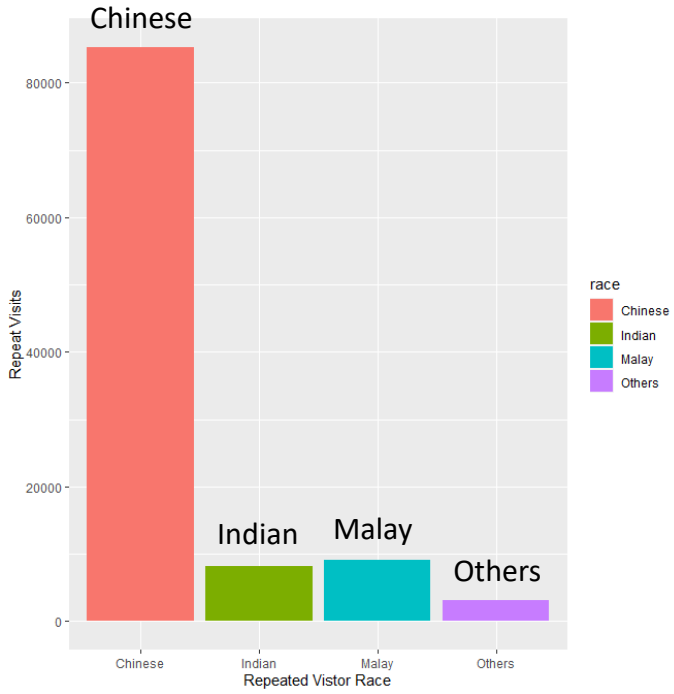
Methodology-Descriptive Statistics

Repeat Visitors

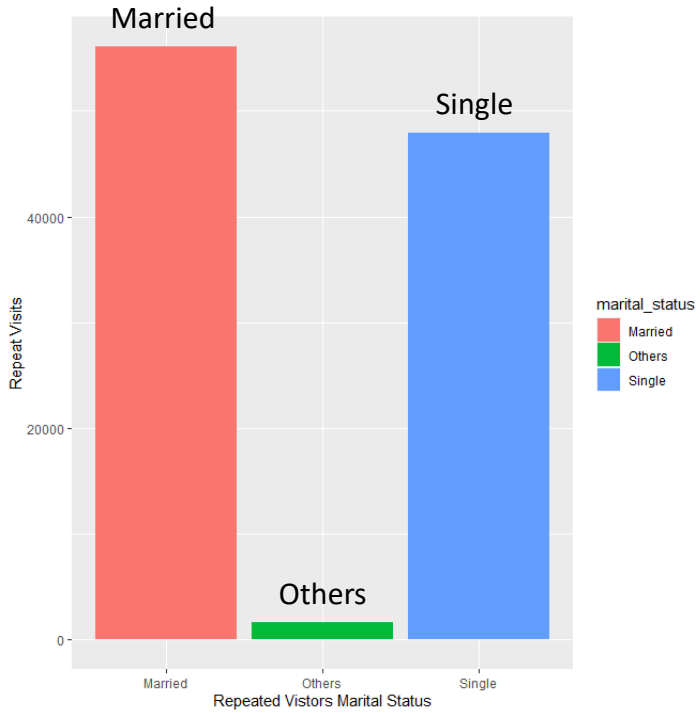
Age Group



Ethnicity



Marital Status



Methodology-Inferential

Change in Visitorship due to Covid-19

H_0: Average Visitors Post Circuit Breaker is greater than or equal to Average Visitors before COVID-19

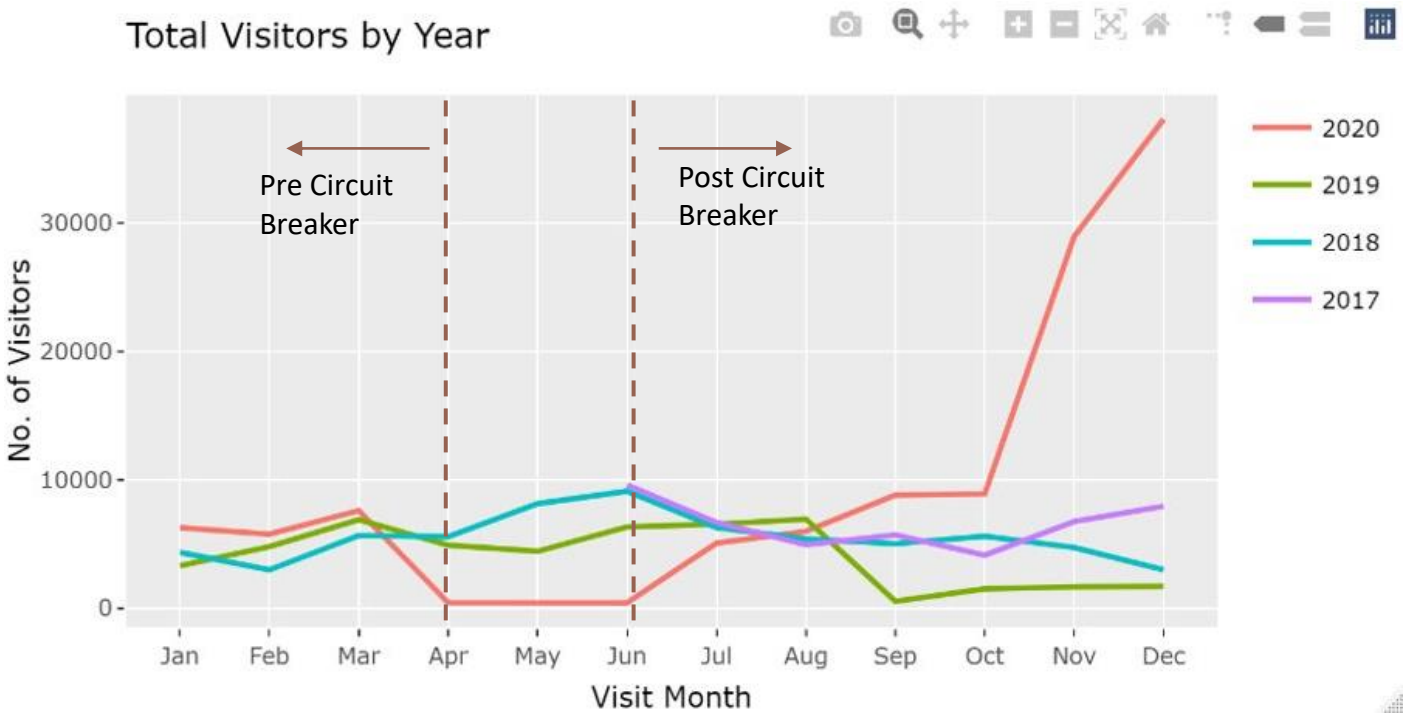
H_1: Average Visitors Post Circuit Breaker is greater than Average Visitors before COVID-19

Circuit Breaker

Result: p-value = 0.9736.

Insufficient evidence to reject the null hypothesis at 90% CI

i.e. H_1 is statistically significant




```
> t.test(mu_x, mu_y,
+        alternative = "less",
+        mu = 0, paired = FALSE, var.equal = FALSE,
+        conf.level = 0.90)

        Welch Two Sample t-test

data:  mu_x and mu_y
t = 2.4801, df = 5.2111, p-value = 0.9731
alternative hypothesis: true difference in means is less than 0
90 percent confidence interval:
     -Inf 2768.145
sample estimates:
mean of x mean of y
 6545.667  4806.333
```

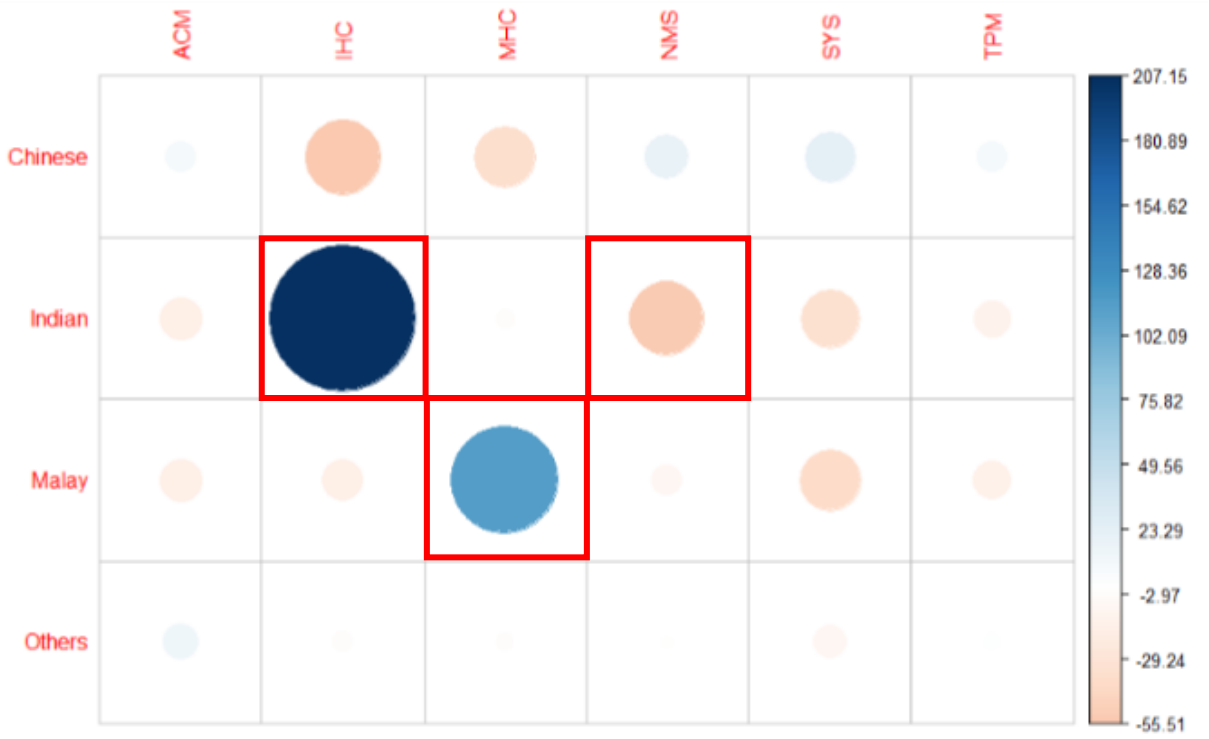
Methodology-Inferential

Chi-Square Test for Demographics Across Different Museum

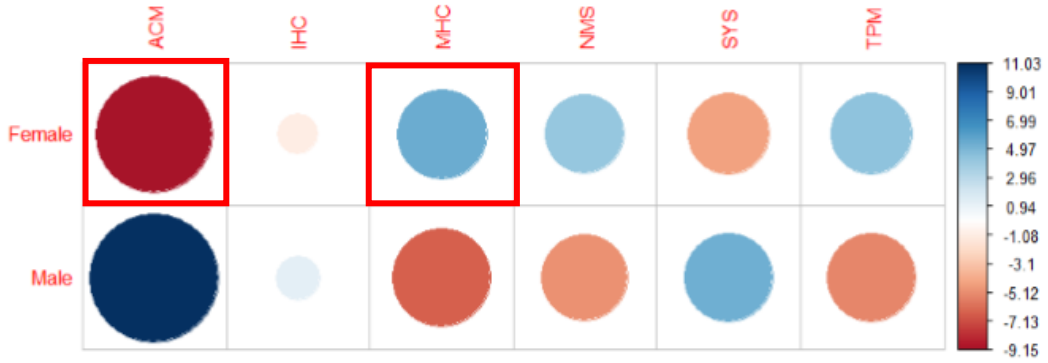
H_0: No association exists between demographics and museum visited

H_1: Association exists between demographics and museum visited

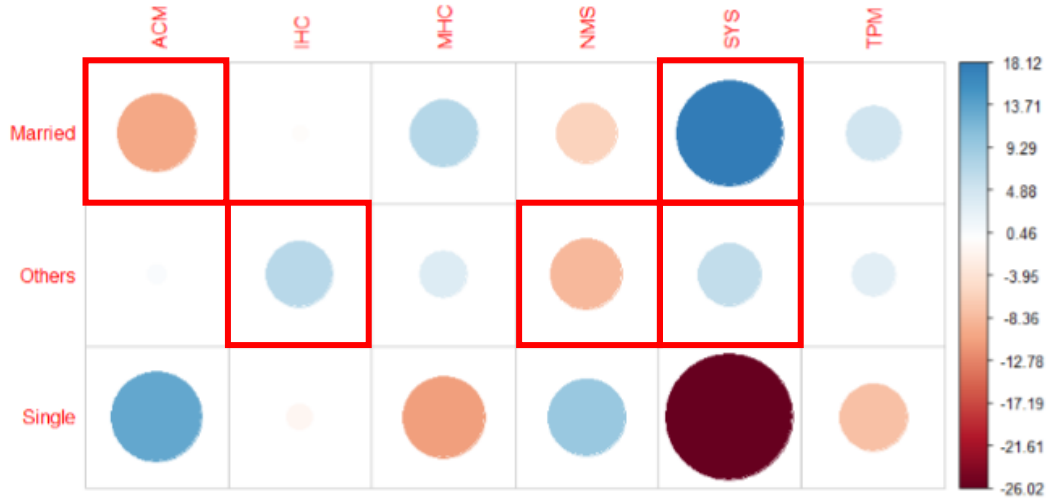
Race



Gender



Martial Status



Methodology-Inferential

Kruskal-Wallis Test for Mean Age

H_0: There is no difference in mean age across museums
H_1: There is a difference in mean age across museums

```
> leveneTest(Age ~ Museum, data = df_age_anova)
Levene's Test for Homogeneity of Variance (center = median)
          Df F value    Pr(>F)
group      5  542.86 < 2.2e-16 ***
      277345
```

```
        Kruskal-Wallis rank sum test

data:  Age by Museum
Kruskal-Wallis chi-squared = 14618, df = 5, p-value < 2.2e-16
```

Pairwise Wilcox test:

	ACM	IHC	MHC	NMS	SYS
IHC	<2e-16	-	-	-	-
MHC	0.26	<2e-16	-	-	-
NMS	<2e-16	<2e-16	<2e-16	-	-
SYS	<2e-16	<2e-16	<2e-16	<2e-16	-
TPM	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16

P value adjustment method: BH

Repeat Visitorship Factor via Multi-Linear Regression

```
Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                      1.74911    0.43508   4.020 5.83e-05 ***
age5  -  14 Years                0.33941    0.43508   0.780  0.43533
age15 - 19 Years                 0.41716    0.43534   0.958  0.33794
age20 - 24 Years                 0.55842    0.43490   1.284  0.19914
age25 - 34 Years                 0.69947    0.43474   1.609  0.10763
age35 - 44 Years                 0.87700    0.43496   2.016  0.04378 *
age45 - 59 Years                 0.97137    0.43493   2.233  0.02553 *
age60 - 69 Years                 1.05222    0.43563   2.415  0.01572 *
age70 - 79 Years                 0.96234    0.43775   2.198  0.02793 *
age80+ Years                     0.89532    0.45500   1.968  0.04911 *
genderMale                       0.11721    0.01881   6.232 4.66e-10 ***
raceIndian                       0.03481    0.03391   1.026  0.30473
raceMalay                       -0.01307    0.03263  -0.401  0.68873
raceOthers                       0.12849    0.05460   2.353  0.01862 *
marital_statusOthers            -0.01964    0.07435  -0.264  0.79163
marital_statusSingle             0.16800    0.02451   6.854 7.28e-12 ***
region_nameEast Region          -0.05730    0.02757  -2.078  0.03772 *
region_nameNorth-East Region    -0.04542    0.02511  -1.809  0.07047 .
region_nameNorth Region         -0.04889    0.03162  -1.546  0.12203
region_nameWest Region          -0.06664    0.02580  -2.583  0.00981 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.
Multiple R-squared:  0.0112
F-statistic: 24.32 on 19 an
```

Adjusted R-squared: 0.01082

Summary: Conclusions

- Inclusivity measured by demographics:
  - Age /Gender/Marital Status/Race
- Accessibility measured by location
- Enjoyability measured by repeat visits

Summary: Recommendations

- Crowd Management in an endemic ‘New Normal’
- Capitalising on local tourism
- Marketing Recommendations
- Accessibility Constraints

Future Work

- Understanding Group Visitors
- Including other variables for regression model
- Using a more relevant benchmark year to assess post-COVID-19 norms


Thank you!