



## **Application User Guide**

**Vast Challenge 2022**

**“Demographics and Relationships  
in the City of Engagement”**

Published by 2<sup>nd</sup> July:

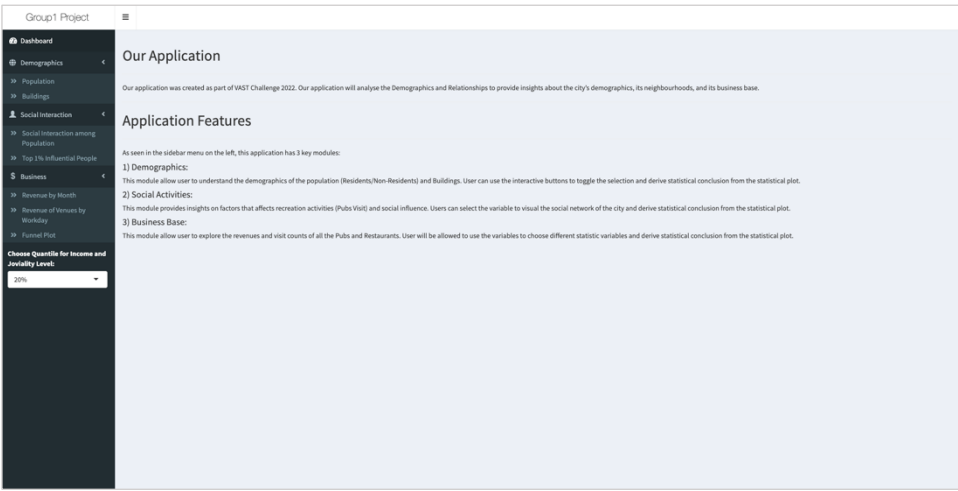
Ong Zhi Rong Jordan

Joyce Chia

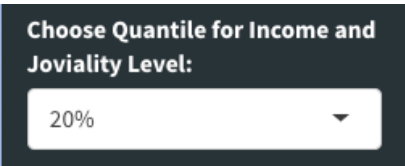
Bomin Kim

### 1. Introduction Page

On this page, there is a short description of the application and an overview of the application features.

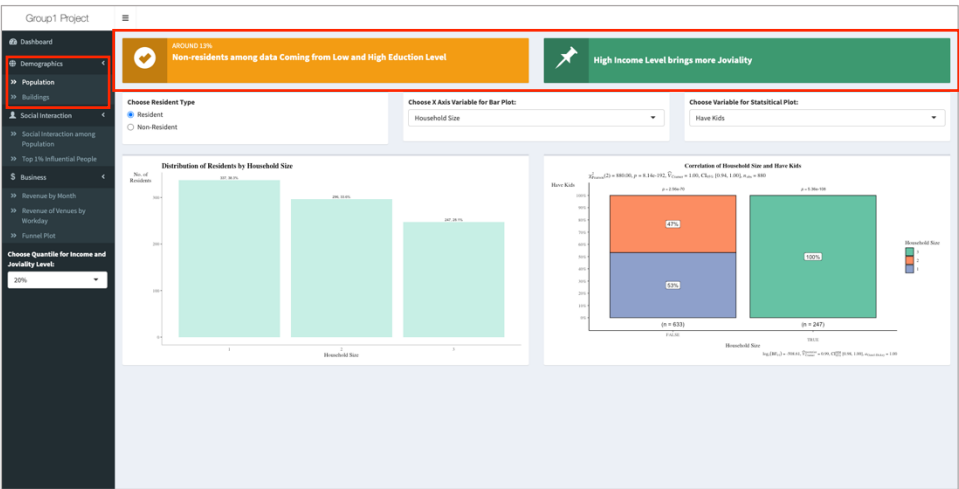


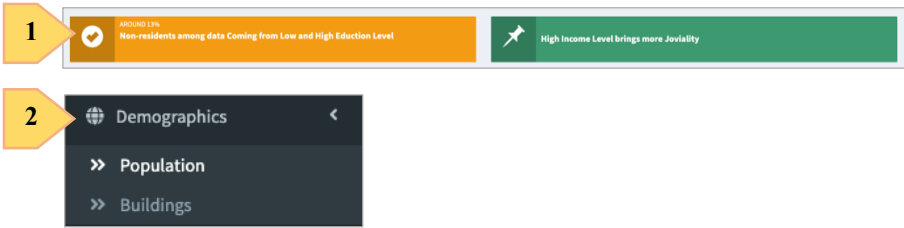
**Note:** Joviality level and Income level are the two most used variables within Analysis for Demographics and Relationships in the City of Engagement. You can always set the desired Quantile (10%-40%) for these two variables before you start the journey to use our application.



### 2. Demographics in City of Engagement

Across each tap and its visualizations, users can understand further details of Demographics in the City of Engagement.



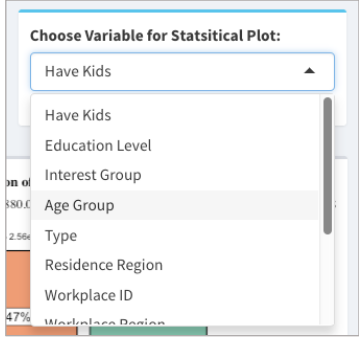
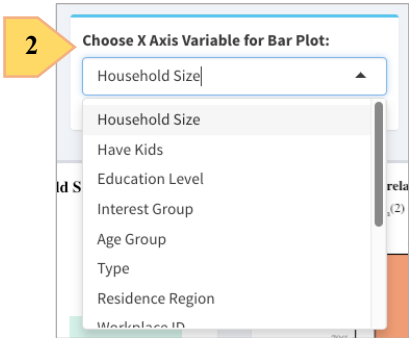
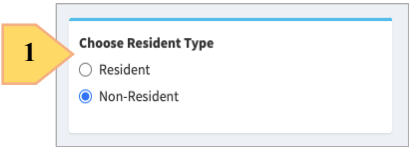
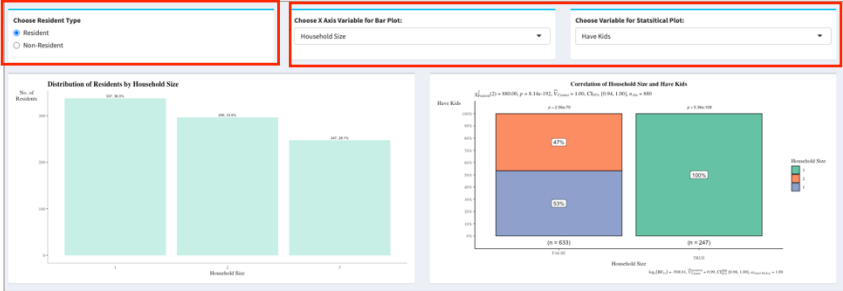


[1] From banners at the very top of the each interface, you can easily identify the main insights of Demographics in the City of Engagement.

[2] Select the categories that you wish to visualize.

### 2-1. Demographics of Population

One static Bar plots and one Statistical plot is use to visualize the Demographics of the Population.

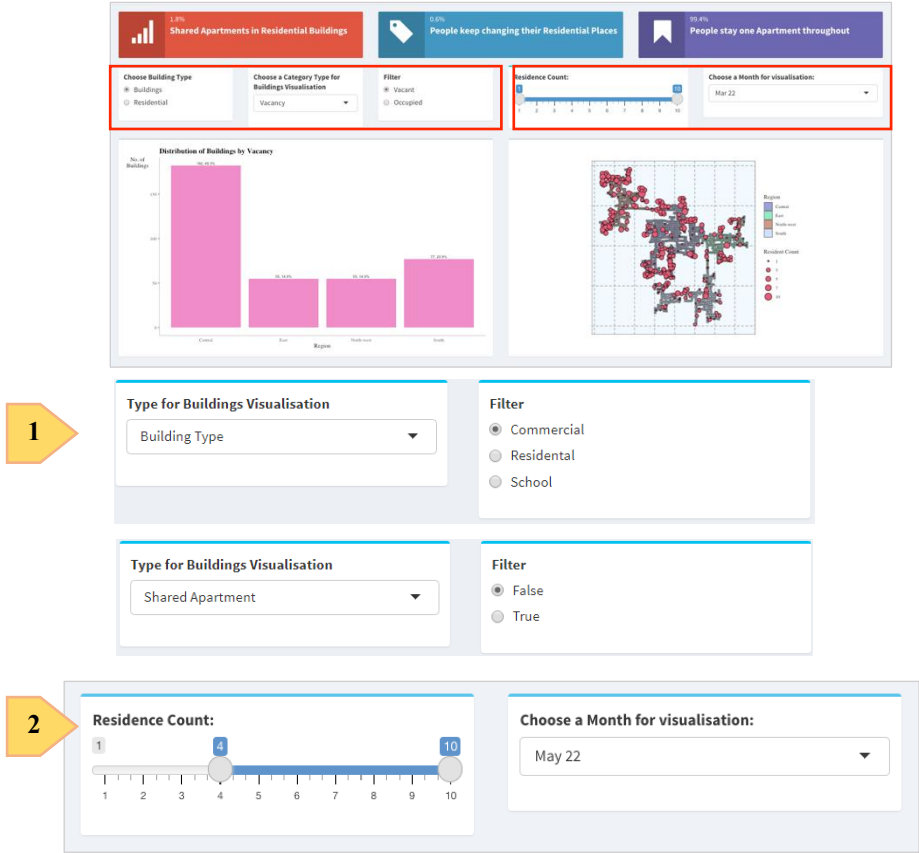


[1] Select Resident type between Resident and Non-Resident, who revealed as staying in the city of Engagement for short period.

[2] Select X-axis variable for both the Bar plot and statistical plot.

### 2-2. Demographics of Buildings

Top half part of the Demographics of the Buildings consists of a static Bar plot and an interactive tmap.



[1] Depending on which building type you select, the Category type for building visualization and Filter will be varied.

[2] Scroll to select the desired range of Residence Count you would like to visualize on the tmap. The range for Month starts from Mar 22 to May 23.

Bottom half part of Demographics of the Buildings consists of interactive datatable based on the brushing of the tmap.

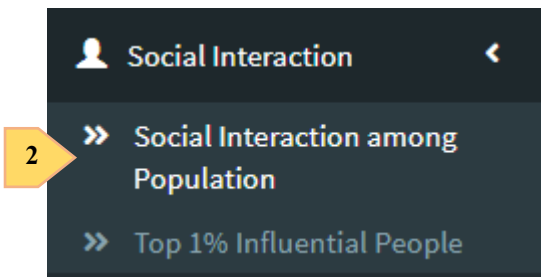
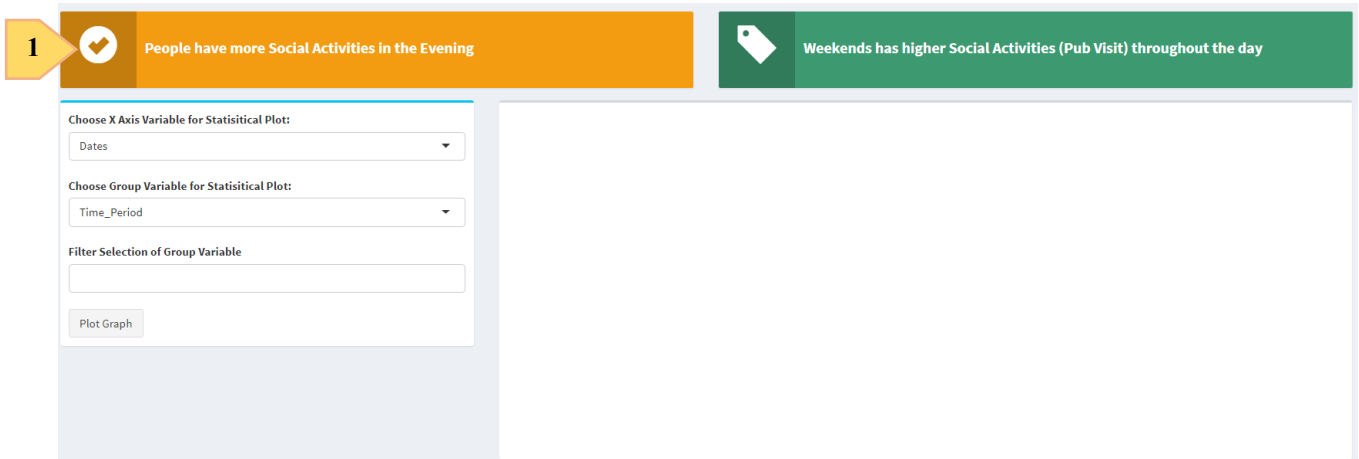
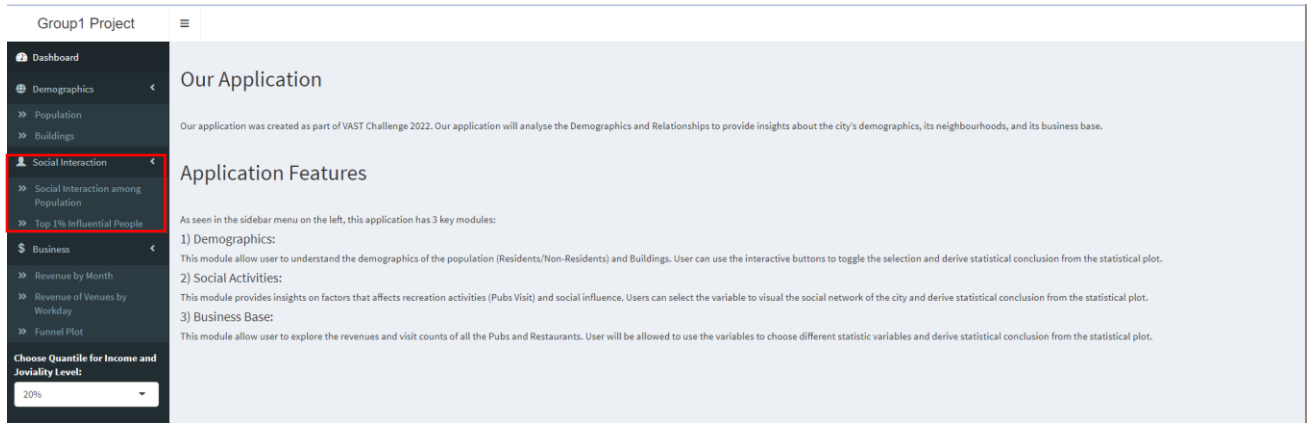
Show 10 entries																	Search:	
	residents	buildingId	Household Size	Have Kids	Age	Education Level	Interest Group	Age Group	Type	Residence Region	Workplace ID	Workplace Region	Income	Rent Apartment	Visit Pub	Joviality	Incc Lev	
1	556	50	3	true	30	High School or College	B	26-30	Resident	Central	1296	Central	40618	False	True	0.626384882	Medi Incor	
2	155	50	2	false	55	Graduate	A	51-55	Resident	Central	1279	Central	168868	False	True	0.176618296	High Incor	
3	143	50	2	false	18	High School or College	A	Below 20	Resident	Central	1280	Central	77028	False	True	0.107753878	Medi Incor	
4	23	1014	3	true	57	Low	B	56-60	Resident	Central	1737	North-west	76545	False	True	0.260211811	Medi Incor	
5	873	1014	2	false	58	Bachelors	J	56-60	Resident	Central	867	Central	62783	False	True	0.282834314	Medi Incor	
6	930	449	1	false	41	High School or College	F	41-45	Resident	Central	434	East	34864	False	True	0.203002706	Low Incor	
7	540	449	3	true	44	Bachelors	C	41-45	Resident	Central	1318	Central	36442	False	True	0.820897579	Low Incor	
8	926	449	1	false	57	High School or College	A	56-60	Resident	Central	877	Central	35258	True	True	0.185717667	Low Incor	
9	582	449	2	false	50	High School or College	D	46-50	Resident	Central	1285	Central	39679	False	True	0.468741471	Medi Incor	
10	752	449	1	false	47	Graduate	E	46-50	Resident	Central	1330	Central	61354	True	True	0.736149674	Medi Incor	

Showing 1 to 10 of 216 entries

Previous 1 2 3 4 5 ... 22 Next

### 3. Social Network Interaction in City of Engagement

Across each tap and its visualizations, users can understand further details of Social Activities in the City of Engagement.

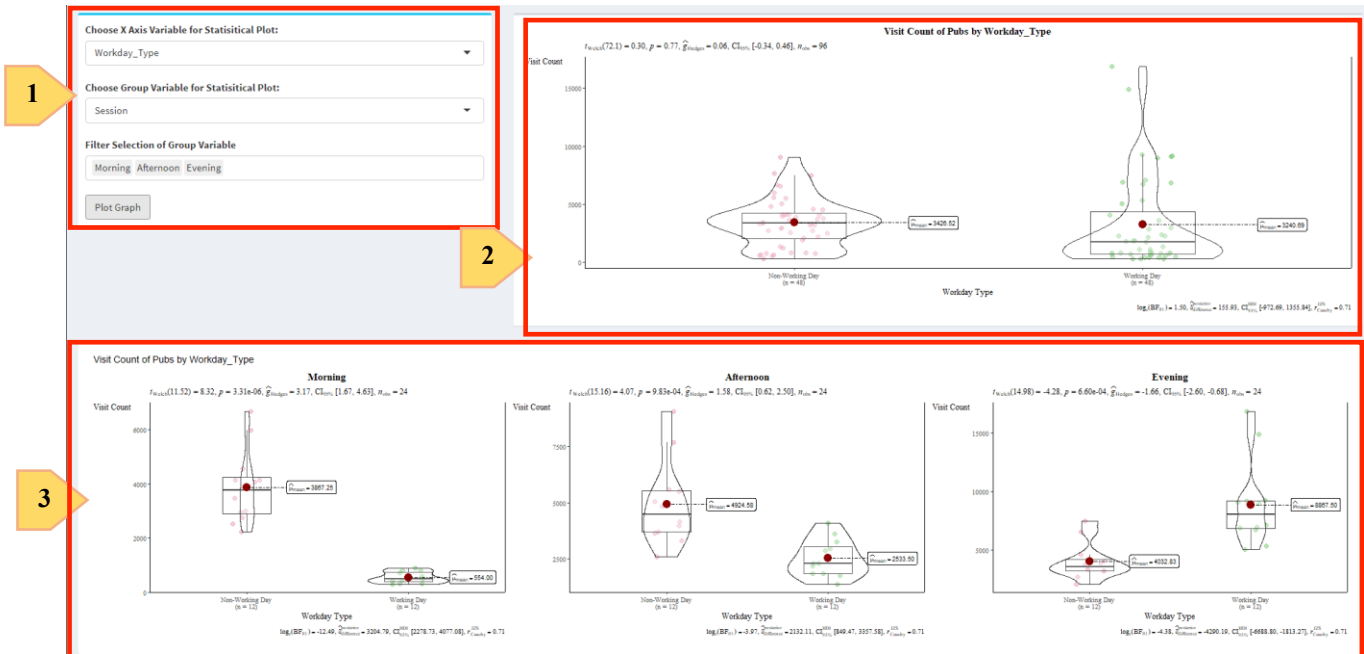


[1] From banners at the very top of each interface, you can easily identify the main insights of Social Activities of the City of Engagement.

[2] Select the categories that you wish to visualize.

### 3-1. Social Network Interaction among the Population

Top half part is consisting of the Social Network Interaction of the Population with Statistical Plot.



[1] Selection of Inputs for the graphs:

[a] **Choose X Axis Variable for Statistical Plot:** User can select any variable to derive statistical conclusion against Pubs Visit Count.

[b] **Group Variable for Statistical Plot** User can select a group variable to be displayed on the grouped statistical plot.

[c] **Choose Filter Type:** User will be able to filter the grouped variable selected at [b] to be displayed on the grouped statistical plot.

[d] **Plot Graph:** Action button to plot the graph after change in either [a], [b] or [c]

[2] **Statistical Plot:** Based on the inputs on [1][a], the statistical plot will either plot a statistical violin boxplot.

[3] **Grouped Statistical Plot:** Based on the inputs on [1][b][c], the statistical plot will either plot a grouped statistical violin boxplot.

### 3-2. Top 1% influential people based on Month and Day

The 2<sup>nd</sup> tab consist of a network graph, ggstatsplot and statistical table to visualize the Top 1% influential participant based on Month and Day type. Users can easily identify the centrality score of each variables and compare each participant's influential impact in the city of Engagement.

The screenshot displays the application's interface with four numbered callouts:

- 1**: Input controls for 'Choose a Month' (Mar 2022), 'Choose a Workday Type' (Working Day), 'Choose a Network Centrality Measure' (Degree Centrality), and 'Choose Variable for Statistical Plot' (Household Size).
- 2**: A network graph titled 'Top 1% influential participant based on Degree Centrality' showing a dense network of nodes and edges.
- 3**: A violin plot titled 'Degree Centrality comparison with Household Size' comparing Degree Centrality across three household size categories: n=137, n=286, and n=247. The plot includes a regression line with the equation  $F_{(2, 567.49)} = 2.44, p = 0.09, \eta^2 = 5.05e-03, CI_{95\%} [0.00, 1.00], \sigma_{\epsilon} = 880$ .
- 4**: A data table showing details for two participants (202 and 255) from the top 1%.

Participant Id	Household Size	Have Kids	Age	Education Level	Interest Group	Age Group	Type	Residence Region	Workplace ID	Workplace Region	Income	Rent Apartment	Visit Pub	Joviality	Income Level	Joviality Level	Degree Centrality	Pub Visit Count
202	1	false	25	High School or College	A	21-25	Resident	East	397	East	41118	False	True	0.985936814	Medium Income	Happy Participant	70	70
255	3	true	59	High School or College	H	56-60	Resident	Central	883	Central	29452	False	True	0.852187763	Low Income	Happy Participant	72	61

[1] **Selection of Inputs for the graphs:**

- [a] **Month:** User will be able to select the month of social interaction they want to visualise on the network graph and statistical plot.
- [b] **Working Type:** User will be able to select the different workday type (Non-Working Days or Working Days) to filter the dataset and visualise on the network graph and statistical plot.
- [c] **Network Centrality Measure:** User will be able to select different centrality measure (Degree, Eigenvector, etc) and visualise on the network graph and statistical plot.
- [d] **Variable for Statistical Plot:** User will be able to select the X axis variable for the statistical plot against the centrality score selected on [c].

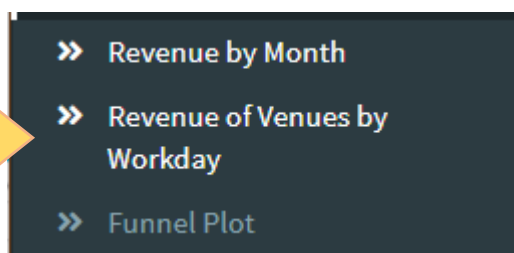
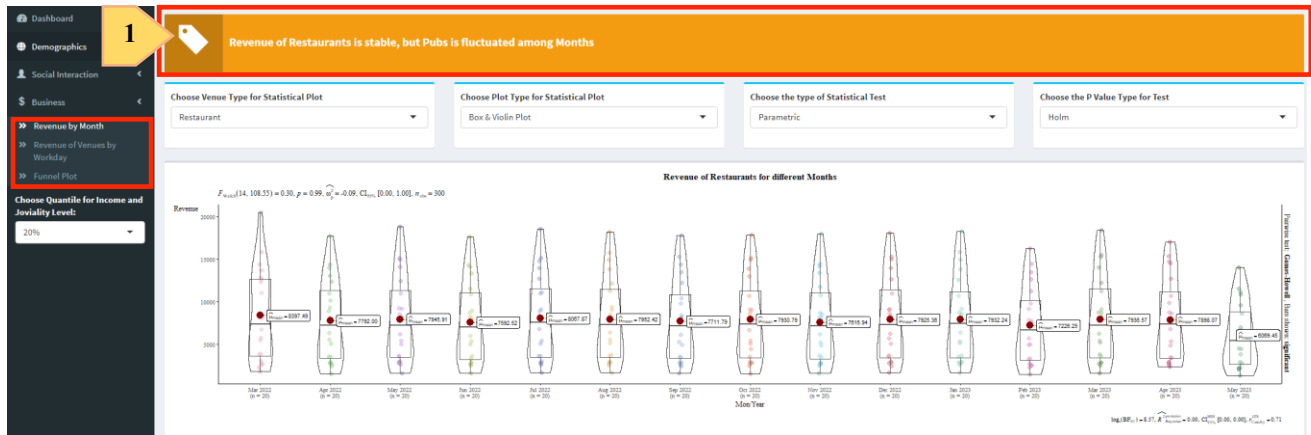
[2] **Network Graph Plot:** Reactive ggraph that display the top 10% nodes based on centrality score and highlight nodes that are the top 1% based on centrality score.

[3] **Statistical Plot:** Based on the inputs on 1[d], the statistical plot will either plot a statistical scatterplot or statistical violin boxplot.

[4] **Interactive Data Table:** Display the details of the participants that belongs to the top 1% from the social network.

### 4. Predominant Business in City of Engagement

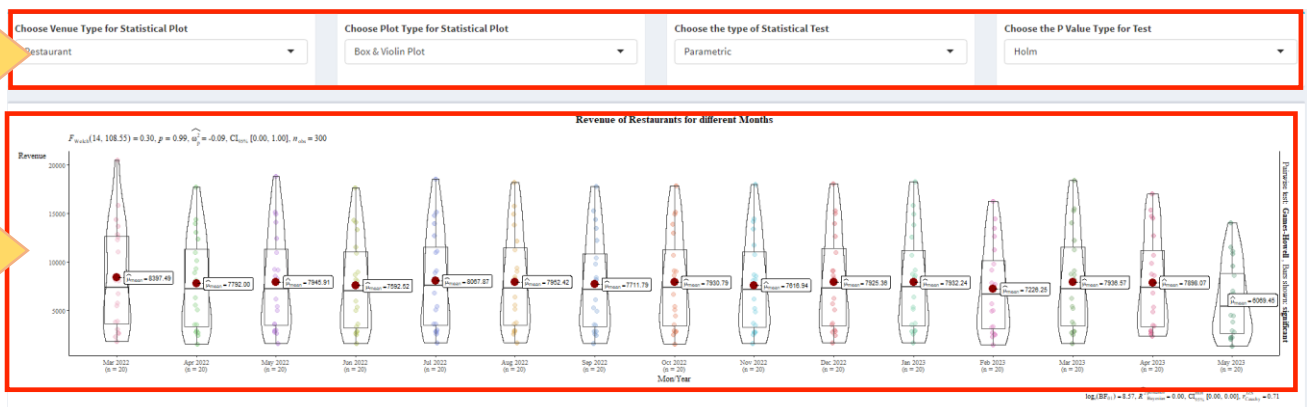
Two of the Predominant Business in the City of Engagement; Restaurant and Pub can be observed in more detail in this section. Users will be able to identify the monthly revenue of two venues and total revenue by two ggstats plots. Tmap will visualize revenue information with geographical distribution.



[1] From banners at the very top of each interface, you can easily identify the main insights of the predominant businesses of the City of Engagement.

[2] Select the categories that you wish to visualize.

#### 4.1 Monthly Revenue of Venue Type



This tab displays the revenue throughout the challenge period for the different Venue Type.

[1] Selection of Inputs for the graphs:



[a] **Choose Venue Type for Statistical Plot:** User can toggle between Restaurants or Pubs to visualize their monthly revenue.

[b] **Choose Plot Type for Statistical Plot:** User can choose either violinbox, violin or box plot to be displayed on the statistical plot.

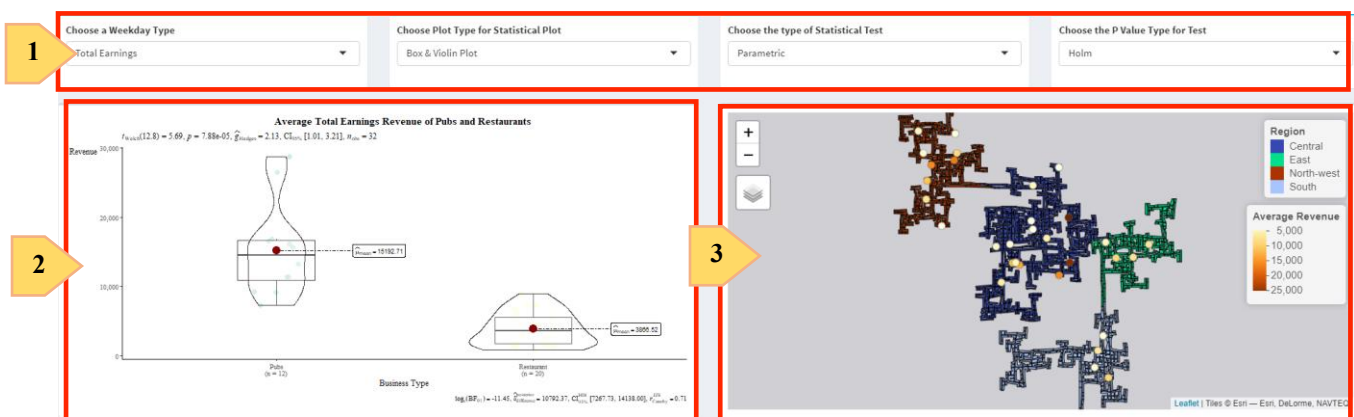
[c] **Choose Statistical Test Type:** User will be able to choose different statistical test type (Parametric, Non-Parametric, etc) to be displayed on the statistical plot.

[d] **Choose P Value for Statistical Plot:** User will be able to choose different P Value type (Holm, Bonferroni, etc) to be displayed on the statistical plot.

[2] **Statistical Plot:** Based on the inputs on [1], the statistical plot will be based on the given inputs and plot the required statistical plot.

## 4.2 Average Revenue of Venue type

This section consist of a statistical plot and a tmap to visualise the average revenue based on the different venue types and the individual venue.



This tab displays the average revenue throughout the challenge period for the different Venue Type.

[1] **Selection of Inputs for the graphs:**

[a] **Choose Workday Type for Statistical Plot:** User can toggle between Total Earnings, Working Day or Non-Working Day to change the statistical plot.

[b] **Choose Plot Type for Statistical Plot:** User can choose either violinbox, violin or box plot to be displayed on the statistical plot.

[c] **Choose Statistical Test Type:** User will be able to choose different statistical test type (Parametric, Non-Parametric, etc) to be displayed on the statistical plot.

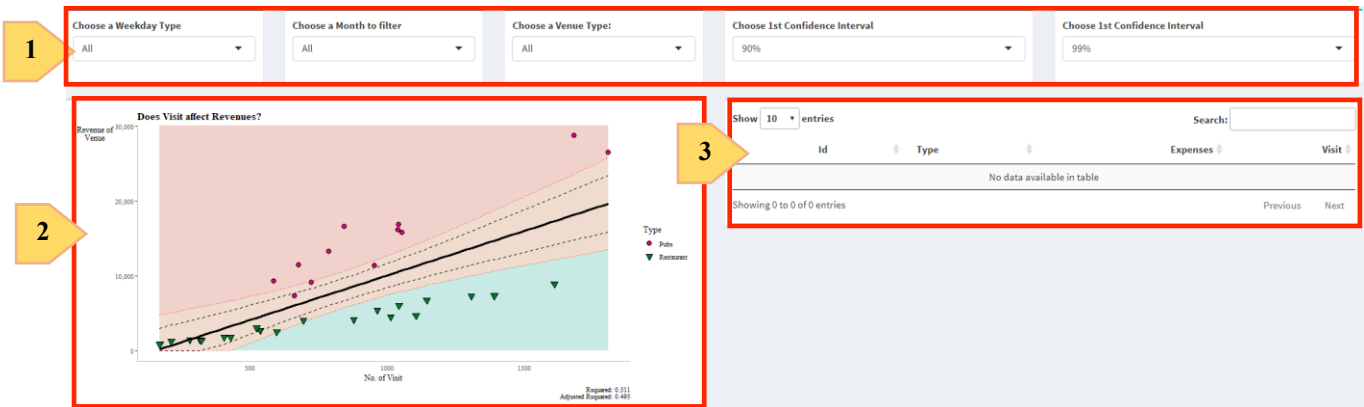
[d] **Choose P Value for Statistical Plot:** User will be able to choose different P Value type (Holm, Bonferroni, etc) to be displayed on the statistical plot.

[2] **Statistical Plot:** Based on the inputs on [1], the statistical plot will based on the given inputs and plot the required statistical plot.

[3] **Tmap:** The user will be able to leverage on the tmap to reveal each venue average revenue throughout the period of the challenge.

### 4.3 Funnel Plot

This section consist of a funnel plot and a datatable to visualise the relationship between Visits and Revenue of the different Venue Type.



#### [1] Selection of Inputs for the graphs:

[a] **Choose Workday Type for Statistical Plot:** User can toggle between All, Working Day or Non-Working Day to change the funnel plot display.

[b] **Choose Month to filter:** User select different months to filter the dataset and visualize the visit and revenue.

[c] **Choose Venue Type:** User will be able to choose either to visualize All Venue Type or just Pubs/Restaurants.

[d] **Choose Confidence Interval:** User will be able to choose different Confidence Interval that will change the visualization of the funnel plot (dotted lines are the confidence interval).

[2] **Funnel Plot:** Based on the inputs on [1], the funnel plot will based on the given inputs and plot the required funnel plot. Rsquared and Adjusted Rsquared will also be displayed at the bottom right of the graph.

[3] **Datatable:** The user will be able to do brushing on the funnel plot and the details of the different venue will be displayed on the datatable.