

How can we better accomodate RCYC's members needs?

Chun Yin Yan

April 1, 2021

Introduction

Topic Description: This reports aims to understand how different groups of RCYC members use RCYC facilities, and what areas RCYC need to work on to better accommodate to the members' needs, based on the data collected on a sample of members in 2017 and 2020.

Research Method: Through the sample data collected from RCYC members, we can better understand the members' habits and needs through statistical inference: improving RCYC services based on statistical information.

Population: The population investigated is all the RCYC members.

Target Audience of Report: RCYC representatives and managers

Data Summary

Data wrangling performed:

- 1 Data filtering to select certain groups of RCYC members, and remove missing values for relevant variables etc.
- 2 Data aggregation to group several similar observations (e.g. usage on sports facilities, money spent on restaurants etc.)
- 3 Data transformation: creating new data (variables) based on the raw data collected for investigation

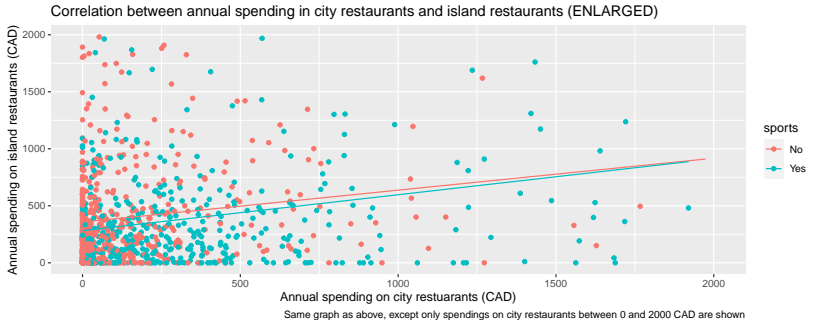
Relevant Variables for each objective:

- 1 Investigation on restaurant spending: “city_dining”, “island_dining” (money spent in each restaurant) etc.
- 2 Investigation on sports facilities usage: “fitness”, “racquets” (whether they used each facility) etc.
- 3 Investigation on spending of dock renters vs non-renters: “city_dining”, “island_dining”, “bar_spending”, “other_spending” (money spent in each category) etc.

Research Question 1

- Objective: Investigate the popularity of RCYC's restaurants in the city of Toronto and restaurants on Toronto Islands
- Question: What is the correlation between the mean yearly amount spent on dining at the RCYC's restaurants in the city of Toronto (mainland) and the mean yearly amount spent on dining at the RCYC's restaurants on the Toronto Islands for fitness and racquets facilities users and non-users?
- Method: A multilinear regression model which shows the correlation between the average annual spending on the restaurants.
- Takeaway: Are the restaurants in the city of Toronto as popular as the restaurants on Toronto Islands? Do members who use fitness and racquets facilities go to RCYC restaurants more?

Research Question 1 - Results



Research Question 1 - Results

A multiple linear regression model for the correlation between mean annual spendings in RCYC city restaurants and islands restaurants between fitness and racquet facilities users and non-users is:

$$y_i = 357.08 + 0.28C - 75.05S + 0.035C \times S,$$

where

- y_i is the estimated annual spending of the individual.
- C is the annual amount spent in city restaurants of the individual.
- S is 1 if the individual uses fitness or racquets facilities, or is 0 otherwise.

We can use this formula to estimate an individual's spending on island restaurants by inputting his/her C and S in the formula.

Research Question 1 - Results

- Finding 1 : Members who do not play sports generally spend less on restaurants
 - Recommendation: Offer promotional discount in restaurants for members who also use sports facilities: this can increase revenue from both sports facilities and restaurants.
- Finding 2: Members generally spend more money on restaurants on the Toronto island. (On average, people who do not spend any money on city restaurants spend over 350 CAD on island restaurants!)
- Recommendation: Offer promotional events and discounts in city restaurants, or gift customers who dine in at island restaurants vouchers which can only be used in city restaurants: this can generate revenue for restaurants in both locations.

Research Question 2

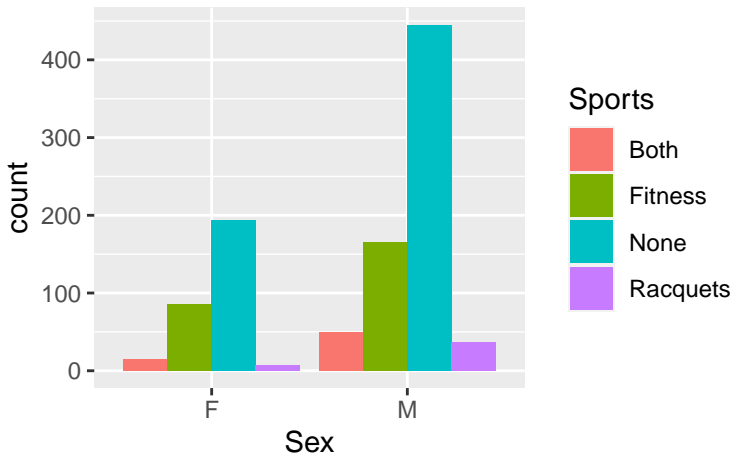
- Objective: Investigate the usage of RCYC's fitness and racquet facilities
- Question: Is the proportion of male members who use fitness and racquets facilities the same as the proportion of female members who use fitness and racquets facilities?
- Method: Two-sample hypothesis testing
- Takeaway: By knowing whether the proportion of male members who use sports facilities is equal to that of female members, we can better accomodate sports facilities users, for example, having more man-size sports equipment.

Research Question 2 - Methods

- Two-sample hypothesis test (Randomization test).
- The hypothesis test is conducted under the assumption that null hypothesis is true. After calculating a p-value which tell us how likely the assumption is true, we can choose whether to reject the null hypothesis. We should state the null hypothesis and alternative hypothesis before calculating the p-value.
- Let p_{male} be the proportion of male members who use fitness and racquets facilities.
- Let p_{female} be the proportion of female members who use fitness and racquets facilities.
- Null hypothesis: $p_{male} = p_{female}$
- Alternative hypothesis: $p_{male} \neq p_{female}$

Research Question 2 - Results

- p-value calculated from hypothesis testing: 0.9426
- Distribution of sports facility usage among RCYC members:



The distribution of sports members do between both sexes are very similar.

Research Question 2 - Results

- Large p-value (0.9426) for null hypothesis shows no evidence against the null hypothesis
- This implies $p_{male} = p_{female}$
- Proportion of members who use sports facilities is 0.35
- No. male members: No. female members in sample = 2.3 : 1
- Recommendation 1: Sports facilities should have enough equipment for both sexes in the right proportion according to the number of female and male members.
- Recommendation 2: Encourage more members to use sports facilities

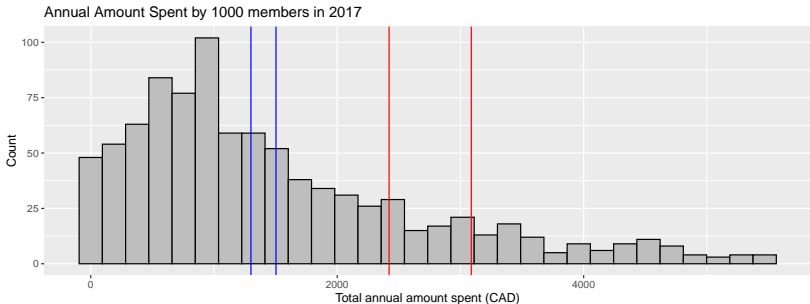
Research Question 3

- Objective: Investigate the differences in amount spent on RCYC facilities between members who rent a dock and members who do not.
- Question: What is a range of plausible values for yearly amount spent on RCYC facilities of members who rent a dock and members who do not rent a dock?
- Method: Bootstrap confidence interval
- Takeaway: Which group should we encourage to spend more money on RCYC facilities?

Research Question 3 - Methods

- Bootstrap confidence interval allows us to estimate the mean spending of dock renters and non renters by giving us a range of plausible values of the mean spending.
- Since the data collected from RCYC members is randomly collected, we can assume that the sample is a representative of all RCYC members.
- Bootstrapping involves random re-sampling with replacement from the original data, then we can create a range of plausible values of mean spending.
- The confidence interval chosen is 95%, meaning that we are 95% confidence that the true means spending of dock renters and non-renters fall within the plausible ranges.

Research Question 3 - Results



This graph shows the annual amount spent by 1000 members in 2017. Note that spendings more than 5500 CAD are excluded for simplicity.

- We are 95% confidence that the mean annual spending of dock renters is between 2421 and 3088 CAD (bounded by red lines).
- We are 95% confidence that the mean annual spending of non-dock renters is between 1299 and 1503 CAD (bounded by blue lines).

Research Question 3 - Results

- Finding 1: Average spending of dock renters is nearly 2 times more than non-dock renters
 - A possible reason might be dock renters spend more time in RCYC facilities than non-dock renters
 - Recommendation: Offer VIP memberships for non-dock renters who cannot afford to rent a dock. VIP memberships offer them discounts to encourage more spending at RCYC.
- Finding 2: Half of the Members who spend more than 4000 CAD per year are non-dock renters (47%)
 - Recommendation: Since dock renters are more likely to have better economic abilities, we should encourage them to spend more money by offering exclusive rewards for dock renters for spending money on RCYC facilities

Limitations

1 Data

- i. We only know whether members use the fitness facility and the racquets facility. If we know the number of hours they use these facilities, we could better accomodate the users.
- ii. The data for restaurant spendings, bar spendings and other spendings are from 2017. The data may not reflect the trends of recent spendings.
- iii. It would be useful to know how long/often members visit RCYC facilities.

2 Research Methods

- i. We assume the data collected to be representative of the whole population (all RCYC members); if not, the accuracy of predictions may decrease.
- ii. The randomization testing (hypothesis test) only tells us the relationship of the variables, but not their casual relationships.

Conclusion and Recommendations

Below are the major findings and recommendations:

- Members spend less on restaurants if they do not use fitness or racquet facilities.
 - Recommendation: Offer discounts in restaurants if member uses RCYC sports facilities.
- Members spend more money on island restaurants than city restaurants
 - Recommendation: Offer promotions in city restaurants.
(e.g. coupons, dinner discounts)
- Around 35% of members use fitness and racquets facilities (sports facilities)
 - Recommendation: Encourage more members to use sports facilities.

Conclusion and Recommendations

- A large proportion of sports facilities users are male.
 - Recommendation: Prepare enough sports equipment for both sexes in right proportions.
- Dock renters spend much more money on RCYC facilities than non-dock renters.
 - Recommendation: Encourage non-dock renters to spend more time at RCYC facilities by offering special memberships/discounts/'Sports Pass'.
- Dock renters do not spend as much money as we expect.
 - Recommendation: Offer exclusive rewards as incentives to spend more money..