

Thanks for taking the time to complete this task.

This project has two objectives: not only to help us see you in action, but also help you see whether the type of work Two Circles does is something you would enjoy.

Gotham Goats is a professional box lacrosse franchise based in Gotham City. They want to build a better understanding of their ticketing customer base. The fictionalized ZIP file you have been given contains 3 sheets - Customer, Events, and Tickets. The *Cust ID* in the Customer sheet is connected to the *Customer ID* in the Tickets sheet, while the *Event ID* in the Tickets sheet is connected to the *Event ID* in the Events sheet.

We would like you to imagine that the Chief Business Officer of Gotham Goats has asked you to pull some key insights out of the data at short notice. Using Tableau please provide: -

* Some analysis of the data quantity and quality
* A maximum of 7 slides, in any format you choose, to identify the key business insights you can identify from the data you have been given

We would like you to share your findings back to us in one week’s time. Please email through your password-protected database analysis to [recruitment@twocircles.com](mailto:recruitment@twocircles.com) in an appropriately secure manner.

Thanks again – we look forward to your thoughts.

The Two Circles Team

**Process – Prepare Data for Analysis**

**Quality and Quantity**

1. **Quality -** used Python - Pandas and Microsoft Excel to check for missing data in the dataset
   1. **Customer Data Sheet – null values were found in multiple columns**
   2. **Should consider collecting all the data fields including Gender, Location info (state, city, zipcode), Company Name (will help us find out their occupation), include a new field column on who do they support as not all customers of Gotham Goats’ are their fan base.**
2. **Quantity –** studied the size of the datasets and its content to see how to structure the analysis

**Findings –**

Tickets and events sheets didn’t have any missing values, while Customer Data had missing values in all the columns including First and Last Names, Address, City, State, Zip Code. Account ID and Country fields showed no missing values.

**Customer Base Analysis**

**In order to answer the following questions, I created four cohorts/segmentations to unpack the customer base. These cohorts are meant to help me unpack the customers of Gotham Goats and find out where are they coming from, who are they, etc**

**Research Strategy -**

It is important to inspect the data to fully understand what kind of information is available for analysis. The initial step to analysing Gotham Goats’ the customer base is to learn who their customers are by creating segmentations or cohorts. (where I grouped cohorts with similar characteristics)

Customer Base Segmentations:

1. **Geographic:** Customer’s Location / Region – *State, City, Zip-Code*
2. **Demographic:** Customer’s Occupation, Gender – *depending on info available*
3. **Behavioural:** Customer Loyalty, Willingness to Purchase – *how many tickets bought? How often?*
4. **Psychological**: Personality Attributes – *chose to look at the field, Opponent*

Some of the Research Questions:

1. **Customers:** Where are they located? What is their age, gender, occupation? Who do they support? What events attended? Was it a game with a specific opponent?
2. **Tickets**: What is the total number of tickets sold every year? Did this number increase or decrease? What is the percent change in ticket sales yearly?
3. **Customers:** What is the total number of customers of Gotham Goats? What is the total number of customers purchased tickets annually? What is the percent change in the number of customers yearly?
4. **Revenue:** What is the total revenue amount per year? Did the amount increase or decrease? What is the year over year growth in revenue from ticket sales?

Customer Base **Geographic & Demographics** –

From the information that was available in the datasets, it was reported that the majority of the customers are male, but it’s important to point out that females come close to 49%. However, there are many missing values in the Gender field, which is skewing the data from being accurate.

Yet, the customer’s geographic location showed that the majority comes from NY, followed by CT, and then Jersey. In addition to this map, there is also a map where the customer base is broken down by Zip Codes and cities.

**Psychological -** When looking at the psychological trait of the customer base, I was hoping to find out to what games do they prefer to attend/watch.

The datasets revealed that the fan base from NY, NJ, and CT all share a preference of watching Gotham Goats’ play against Boston Bears. Games against Detroit Wheels were popular among New Yorkers and New Jersey customers while Montreal Concords are popular among Connecticut peeps.

**However,** the psychological traits found in this analysis does not specify whether customers included in the datasets support Gotham Goats or the opponent team.

A more in-depth analysis would need to be done in order to find out whether they are Gotham Goats supporters, which could be done by looking at the ticket sales per Account\_ID and grouped by the opponent field. This will show individual accounts frequency of purchasing tickets, attending games vs opponents

**Behavioural Traits of Customer Base**

Behavioural Trait looked at customer loyalty and willingness to buy. In 2015, Gotham Goats faced a significant drop of 19% in ticket sales. Yet, in 2016 and 2018 the total number of tickets sold to the customer increased annually representing strong revenue retention. However, year over year growth of ticket sales in 2018 was reported lower than in 2017.

**Customer & Revenue Retention**

1. The total revenue amount earned from the ticket sales fell by almost 15% in 2015, but the amount increased by 31% in 2016, 44% in 2017, and 49% in 2018.
2. In 2018, price per ticket reported as highest
3. Lowest Max price per ticket was in 2015, the same year as the number of tickets sold decreased.
4. Gotham Goats’ show to have a strong revenue retention when it comes to selling the tickets to their games.
5. The revenue amount rose by 49% in 2018, but the total number of customers who purchased game tickets in the same year reported to be at a lower rate than in 2017.
6. The min price per ticket stayed the same, but max kept fluctuating between high 50s and low 60 but jumped to $73 in 2018. Makes me think that their customer retention is weak.