Analysing Restaurant Ratings and Success Factors - A Data Science Approach

Capstone Project by Mehak Sethi
(Springboard DSC+ - April 2020 Cohort)

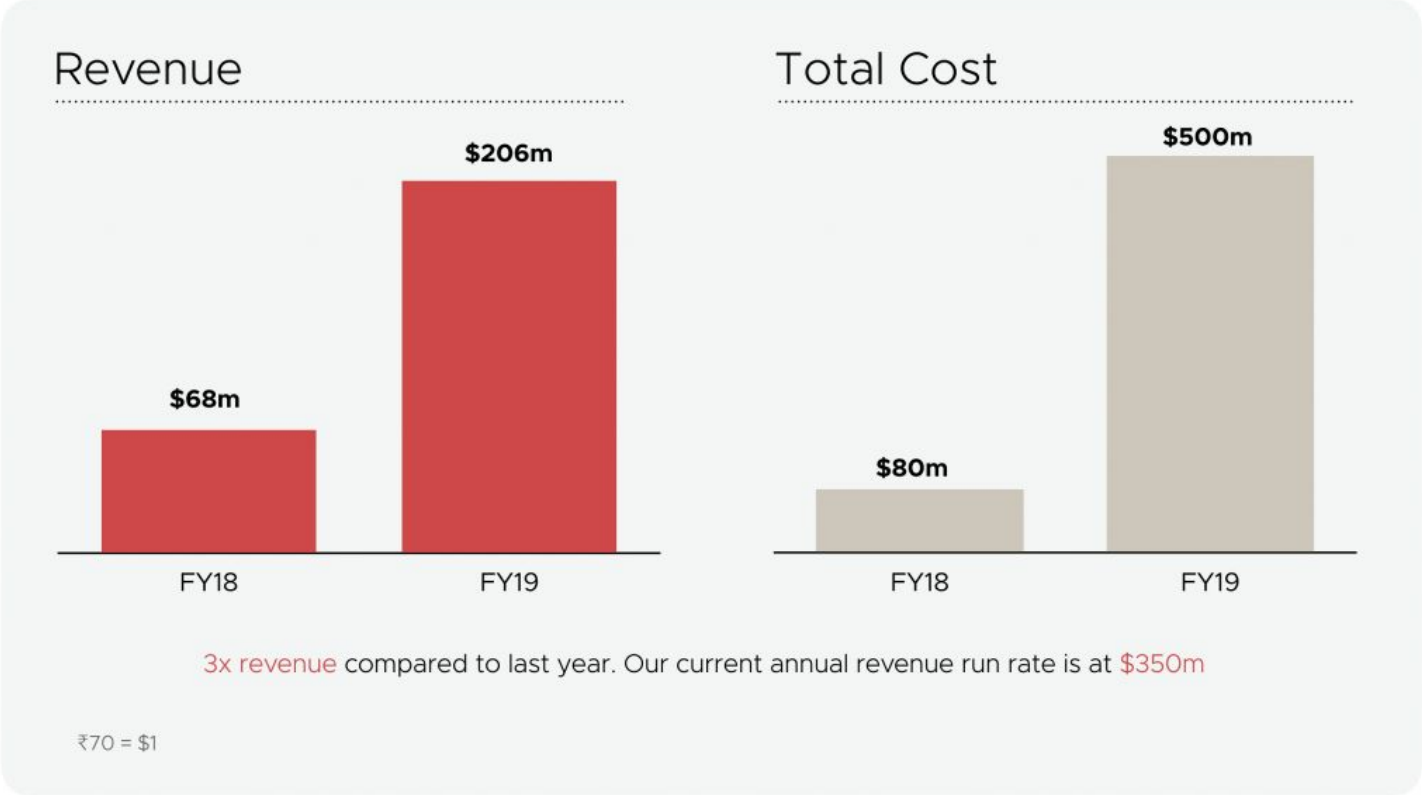
About Zomato

(Business Context)

Zomato is one of India's favourite food aggregators. It was established 11 years ago and today it operates in 24 countries and 10000+ cities globally. The company has a simplistic vision of BETTER FOOD FOR MORE PEOPLE.

It started out with providing restaurant listings in an app, it was much like the Yellow Pages of restaurants. Fast forward to the present day, Zomato has branched out into food delivery at deep discounts, providing high quality raw materials to partner restaurants (HyperPure by Zomato) to restaurant promotion via advertising, delivery of groceries to organising India's biggest food fest Zomaland to providing meals to the underprivileged.

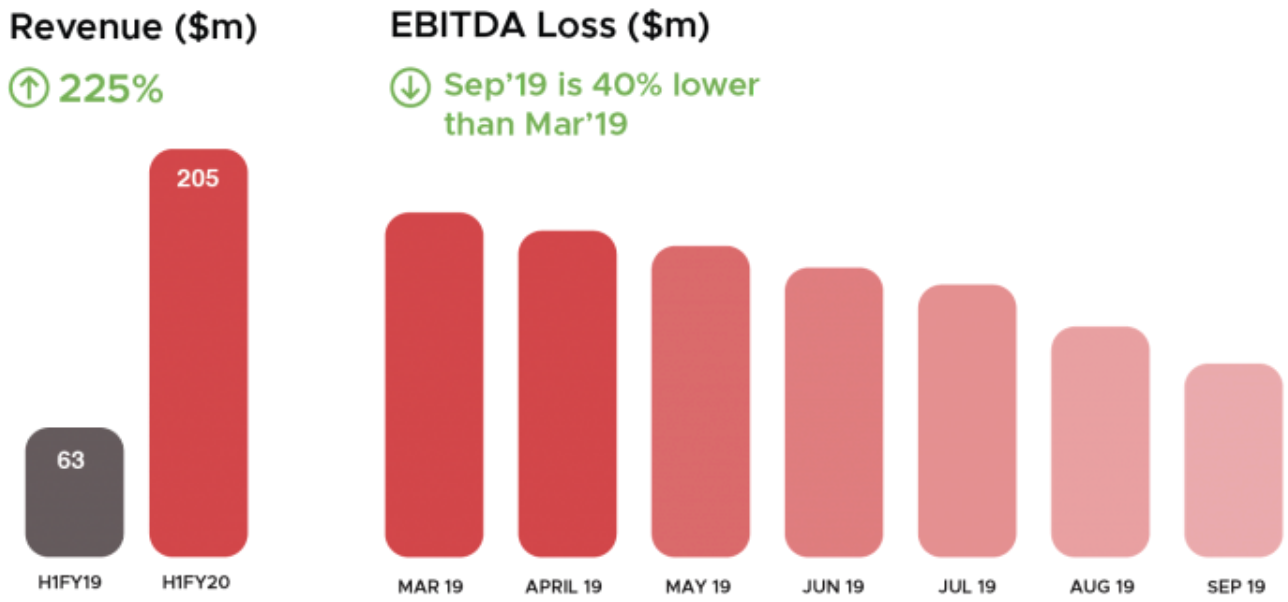
Their revenues have grown multifold y-o-y and so has the cost (snapshot below from the Annual Report 2019).



Delivery		Dining Out		Sustainability	
REVENUE		REVENUE		REVENUE	
FY19	$155m	FY19	$49m	FY19	$2m
FY18	$38m	FY18	$30m	FY18	$0

(Source: Annual Report 2019)
The revenue from the food delivery operations, dining out and sustainability have increased in the last year.

Increasing Revenues and Increasing Losses



(Source: Half Year Report H1 2020)

Despite being one of India's top food aggregators, Zomato is also in heavy losses and much of it is because of high expenses on delivery operations. However it is the market leader in that segment. In recent years, it has started loyalty membership called Zomato Gold for its users wherein they can pay an annual fee and enjoy free food and beverages at partner restaurants. Another model of income is advertising. Restaurants can pay a premium to Zomato to be displayed on top of the listings and can also advertise their restaurant on the app.

In the middle of difficulty, lies the opportunity!

Business Problem

Zomato has been in the business for 11 years and still is not profitable. Through Zomato Gold and Advertising, it is clear that it is trying hard to generate revenues but currently it is not that effective as seen in the numbers.
However, it has 11 years of data of each user's dineline, of majority restaurants, in almost all parts of the country and even abroad. It can leverage this gold mine (data) and get to know what makes a restaurant a best performer, i.e. Restaurant Ratings can be predicted. It can then use these predictions to offer consulting services to upcoming restaurants and troubled outlets which could leverage the insights to turn their business around. That could be a win-win for the restaurant and Zomato and both would earn more revenue on it. So through this project, I aim to analyze what factors contribute to Restaurant Ratings and predict them.

Data Set Link :
https://drive.google.com/file/d/1N1wyGSK26uF6CbA9y509UPaSaz7S4xoI/view?usp=sharing
(Dataset sourced from Kaggle)

Scope of Solution Space:
I want to make use of Exploratory Data Analysis techniques, clustering, decision trees and linear regression and some advanced machine learning techniques (if required) to:

- Drill insights about what is present in the data
- Identify segments of restaurants based on various factors
- Attempt to visualise a decision tree to easily uncover patterns in the data
- Make predictions of Restaurant Ratings using Linear Regression (or advanced ML Models)

Constraints:
The data regarding retail consumers (demographic, dineline etc) is not available. Ordering Data is not available.

Stakeholders:
Since this is an academic project, the key stakeholders are:

- Me
- My Mentor (Mr. Navaneesh Gangala)
- Other members of Springboard India team

Timeframe:
The budget timeframe is about a month (subject to course deadlines).

Success Criteria:
I aim to predict Restaurant Ratings with decent values of Explained Variance Percentage, Adjusted R- Square and other accuracy metrics as per the course guidelines.