# IF-7300

*by* A a

**Submission date:** 30-Oct-2023 06:27AM (UTC-0500)
**Submission ID:** 2163963643
**File name:** IF-7300.txt (9.59K)
**Word count:** 1517
**Character count:** 8192

1. Introduction

Business analytics (BA) refers to the refinement of historical and present business data through the use of advanced techniques such as descriptive statistics, machine learning and more. BA includes data collection, sequence identification, predictive analytics and visualisation, which helps businesses to make data-driven decisions (Ahmad et al., 2022). Divvy is a Chicago-based bike-share system, operating through more than 100 stations (Divvy, 2023a). The primary purpose of this study is to apply the concept of programming in business analytics to collect, preprocess and analyse bike sharing data of DIVVY for 2017 (Quarter 1 and 2).

2. Methodology

2.1 Data Description

Figure 1: Overview of the variables

Dataset is based on the historical trip data of Divvy for 2017, quarter 1 and quarter 2. Each trip of Divvy includes trip start day and time, end day and time, start station and end station, rider type, user type and trip duration (Divvy, 2023b). From Figure 1, it can be found that the variables available in the quarterly dataset of Divvy are trip_id, start_time, end_time, bikeid, trip_duration, from_station_name, from station_id, to_station_id, to_station_name, user type and gender.

Figure 2: Variable type

Within the dataset, there are integer, string and float values presented. Variables like bikeid, trip duration and station IDs are integer types. On the other hand, variables like user type (three classes: subscriber, customers and dependent) and gender (two classes: male and female) are categorical variables.

## 2.2 Data Preprocessing

Figure 3: Checking for null values in the datasets

Data preprocessing and cleaning is a crucial step in business analytics as the elimination of null and redundant values increases reliability of the data analysis. For the elimination of null values from the dataset, 'isnull().sum()' function has been used.

Figure 4: Data Cleaning and preprocessing

Data dropping is an important method of data cleaning. Through the application of data dropping technique, the elimination of rows with missing start and end data has been performed. Trips that are less than 1 minute have been dropped to ensure the fact of potential false starts and re-docking the bikes for security checks. Duplicate values were dropped from each variable by using the drop_duplicates () function [Refer to Figure 4]. The elimination of the null and duplicate values from each of the columns has ensured increased data reliability and accuracy for data visualisation and analysis.

Figure 5: Shape and Size of the dataset after cleaning

The elimination of the null and redundant values has ensured the development of a structured dataset for further analysis. The shape of Q1 and Q2 datasets after the data pre-processing and cleaning operation are (431691,12) and (1119803,12) [Refer to Figure 5].

# 3. Results and Analysis

## 3.1 Descriptive Statistics

Figure 6: Descriptive Statistics for trip duration in Quarter 1 and Quarter 2

As per the viewpoint of Ukam et al. (2023), trip duration fundamentally depends on different factors including non-trip stops, type of trips, recurrent congestion index, signal delay and more. In Quarters 1 and 2, Divvy bike trips have exhibited diverse durations [Refer to Figure 6]. In 2017 (Q1), the average trip duration for Divvy was 742.64 seconds (about 12 minutes) with a standard deviation of 1514.94 seconds, indicating a wide range of trip lengths for Divvy. On the other hand, trip duration in Q2 increased to 1015.78 seconds (around 17 minutes). This signifies that the organisation has increased the trip duration by serving a wide customer base. On the contrary, the high standard deviation in trip duration both in Q1 and Q2 demanded the optimisation of bike-sharing systems for different user preferences.

Figure 7: Descriptive Statistics for Station Capacity

Descriptive statistics for Station capacity have revealed detailed insights related to the distribution of available docks for Divvy. The mean station capacity of 17.18 docks with a standard deviation of 6.10, indicate that there was a moderate level of variability in station sizes across their 582 stations. Most of the Divvy bike stations offered 15 to 19 docks, while the maximum station capacity was 51 docks. This has showcased that stations of Divvy bike have a large infrastructure.

## 3.2 Revenue generated from each user types

Figure 8: Revenue generation from each user type

Revenue of Divvy bikes is significantly dependent on the type of users. The three main customer categories for Divvy are Annual Member, Single Ride and Day Pass. From figure 8, it can be observed that revenue generated from subscribers (annual members) increased by $26476665264 to $66804899472 from 2017 (Q1) to 2017 (Q2). In fact, revenue generated from annual members was almost 80% of total revenue in 2017. The primary reason for the enhancement of market revenue from subscribers was the increased brand loyalty and high conversion rate from customers to subscribers. On the other hand, Divvy generated a revenue of $113522465 in Q1 which increased to $7783527000 in Q2. This signifies enhancement in the number of daily customer base in 2017, Q2, due to extensive marketing campaigns and high service quality.

3.3 Proportion of subscribers versus Customers at Divvy

Figure 9: Proportion of subscribers versus Customers at Divvy

The overall proportionality of subscribers versus customers in the Divvy bike-sharing system is a key metric for the identification of user demographics as well as system engagement. From figure 9, it can be observed that 73.6% of users of Divvy bikes were annual members (subscribers) and 26.4% of the users were one-time customers. The existence of member-only perks like “unlimited 45-minute rides on classic bikes”, “ebike prices and scooters”, “Bike Angels”, “Community Events” and reasonable pricing starting from $130.90/year has helped Divvy to convert one-time customers to subscribers (Divvybikes, 2023). Hence, it can be stated that a higher proportion of subscribers indicates loyal user bases and retention of diverse user profiles for sustainable growth.

3.4 User type distribution of the organisation

Figure 10: User Type distribution for Divvy

Figure 10: Visualisation of User Type distribution for Divvy

According to the data of Q1, 390,292 annual members make up the majority of the user base of Divvy, followed by 41,395 customers and very few dependents [Refer to Figures 10 and 11]. Customers are greatly outnumbered by subscribers, indicating a robust base of yearly members. In fact, in Q2, the number of subscribers increased significantly to 844,047, and the number of customers to 275,756. This spike indicates that the subscriber base of Divvy Bikes has grown significantly, possibly as a result of successful retention tactics, marketing campaigns or growing interest in yearly memberships.

3.5 Station Capacity

Figure 11: Station capacity at Divvy

The station capacity for Divvy bikes has significantly varied across different stations. As per the viewpoint of He, Ma and Jin (2021) and Hurtubia, Mora and Moreno (2021), within the context of the bike-sharing system mean station capacity indicates the availability of docks per station. The value of mean station capacity for Divvy in 2017 was 17.18, which signifies that on average 17.18 docks were available for bikes in each of the stations.

3.6 Customer Segmentation of Divvy

Figure 12: customer segmentation at Divvy in 2017, Q1

From figure 12, it can be found that 80% of the users of Divvy bikes were male and 205 of the customers were female in 2017 (Q1). This signifies the popularity of bike-sharing services among male customers in Chicago.

Figure 13: customer segmentation at Divvy in 2017, Q2

In Quarter 2, the number of female users has further decreased. A possible reason for that is the non-existence of bikes designed especially for women. The number of female users increased from 81663 to 217121 in 2017, Q2 [Refer to Figure 13].

3.7 Evaluation of trip count for quarter 1 and quarter 2

Figure 14: Trip duration at Divvy in 2017, Q1

The trip duration for Divvy significantly dropped in March 2017. Decrement in perceived travel time and degradation in perceived travel experience creates a negative impact on the trip duration (Pourhashem et al., 2023). These can be reasons for the sharp decline in trip duration for Divvy bikes in 2017, Q1.

Figure 15: Trip duration at Divvy in 2017, Q2

Trip duration has significantly increased in the second quarter of 2017. Effective utilisation of station capacity and docs along with enhancement of the number of subscribers has helped the organisation to increase the trip duration.

## 4. Recommendation

Divvy needs to adjust station capacities based on demand patterns and customer segmentation in order to ensure that stations align with user needs.

The organisation needs to introduce targeted marketing campaigns, loyalty programs and incentives for subscribers to enhance the number of subscribers

Divvy needs to optimise pricing structures and service offerings along with the placement of stations to increase the trip duration.

Divvy needs to introduce bike designs particularly for women to attract more female users.

## Conclusion

Based on the above discussion it can be concluded that effective utilisation of docks has helped Divvy in increasing their trip duration. In fact, the capability of converting one-time customers to annual members has increased their revenue in 2017. Moderate diversity in station capabilities indicates a well-balanced network. Divvy needs to strategically promote subscription-based usage given that it saw a significant rise in customers in the second quarter.

IF-7300

ORIGINALITY REPORT

| 4% | 2% | 2% | 3% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Submitted to University of North Texas<br>Student Paper | 1% |
| 2 | Submitted to Monash University<br>Student Paper | 1% |
| 3 | Submitted to Southern New Hampshire University - Continuing Education<br>Student Paper | 1% |
| 4 | Mohammed A. Mahyoub, Ravi R. Yadav, Kacie Doughetry, Ajit Shukla. "A Machine Learning Model Integrated with the Clinical Workflow Detects Sepsis Early with High Sensitivity and Specificity", Cold Spring Harbor Laboratory, 2023<br>Publication | 1% |
| 5 | www.zbw.eu<br>Internet Source | 1% |

| | | | |
|---|---|---|---|
| Exclude quotes | Off | Exclude matches | Off |
| Exclude bibliography | On | | |