

Minor Project Report
On
Metro Bike Share Trip Data

by

Neelotpai Singh Gour

Registration number: 11611919

Roll No: A11

Course Code: INT217

Section: KEM41



School of Computer Science and Engineering
Lovely Professional University G.T Road Jalandhar
Phagwara, Punjab (India) 144401

DECLARATION STATEMENT

I hereby declare that the project entitled “**Metro Bike Share Trip Data**” submitted at Lovely Professional University, Phagwara, Punjab is an authentic work and has not been submitted elsewhere.

I understand that the work presented herewith is in direct compliance with Lovely Professional University’s Policy on plagiarism, intellectual property rights, and highest standards of moral and ethical conduct. Therefore, to the best of my knowledge, the content of this case study represents authentic and honest effort conducted, in its entirety, by me. I am fully responsible for the contents of my case study report.

Signature of Candidate

Neelotpal Singh Gour

Reg No: 11611919

CERTIFICATE

This is to certify that the declaration statement made by me is correct to the best of my and belief. They have completed this Project under my guidance and supervision. The present work is the result of their original investigation, effort and study. No part of the work has ever been submitted for any other degree at any University. The Project is fit for the submission and partial fulfillment of the conditions for the award of B. Tech degree in Computer Science Engineering from Lovely Professional University, Phagwara.

Signature and Name of the Mentor: Robin Kumar Mathur

Designation: Assistant Professor

School of Computer Science and Engineering,
Lovely Professional University, Phagwara, Punjab.

Date:18/11/2018

ACKNOWLEDGEMENTS

We would first like to thank friend whose kind and constant support during the project was our biggest strength, and to whom this project has been dedicated.

We would like to my sincere gratitude to the **faculty**. To give us information of such a great industry and his inspiration proved to be the path finder.

Table of Content

S.No.	Title
1.	Introduction
2.	Scope of Analysis
3.	Existing System
4.	Source of Dataset
5.	Analysis on Dataset
6.	Future Scope
7.	References
8.	Bibliography

Introduction

Bike sharing is a service which provides available bikes as a shared use for individuals on a short-term basis, either free or at a reasonable price. A metro bike sharing system allows users to rent a bike from one station and return it at any other station within the system. Metro bike share systems have been deployed in various cities around the world since the second half of 20th century and become more popular in recent years. These systems provide access to bicycles for short-distance trips as an alternative to private vehicles or motorized public transport such as bus or subway in an urban area. In addition, they help reduce the traffic congestion, air pollution and noise. Moreover, they have been considered to solve the “last mile” problem. Finally, they help bridge the gap between existing transportation modes such as subways and bus systems and connect users to public transit networks. Besides the benefits mentioned above, systems face many problems, one of which is the availability imbalance. Since movements of customers are highly dynamic, the bike usage is non-stationary, changing markedly with time and location. Therefore, some stations may be short of available bikes for rent while some are full and do not have enough docks for returned bikes. A general approach to solve this problem is that the system should monitor and redistribute bikes between stations frequently using trucks or bike-trailers.

Of all the 21st century smart-city data we have, virtually none of it directly answers one of the most important questions about a city: Where is people going?

If you think about it, we regularly count where people are at a certain point. There's data for what bus stops and train stations people use. There's data for vehicle traffic throughout the city, taken at specific points. Sometimes we do pedestrian and bicycle counts, which are laborious. In short, we have a lot of Point As, but virtually no Point Bs.

Scope of Analysis

Unbalanced stations

Bike sharing programs usually have a problem of unbalanced stations where the number of trips from these stations is higher than the number of trips to these stations (or vice versa). Because of this issue there is a need to transfer bicycles using trucks between stations.

Not uniform usage of bicycles

Some stations are very popular with many rents, while some have only few rents. Because of that in general bicycles at popular stations tend to be used significantly more often than bicycles at not popular stations. A not uniform usage of bicycles leads to a need of bringing heavily used bicycles often to a workshop, while there are some bicycles almost new and used only a few times.

The goal is to analyses the data and see if there is a possibility to suggest bicycle transfers in a way to balance bicycle usage.

Source of Dataset

Kaggle-

<https://www.kaggle.com/cityofLA/los-angeles-metro-bike-share-trip-data>

Analysis on Dataset

1. Calculate total no of rides till now

General Description- In this analysis we have calculated the total no of rides, which is displayed on dashboard. We use count function to calculate the total number of rides.

Function- COUNT

Formula- COUNT(Trip)

Analysis Result- 132,417

Visualization-

Total Rides	132,305
-------------	---------

2. Find out the busiest station

General Description- In this analysis we have analyzed the data to find the busiest stations, busiest means where bike incoming and bike outgoing is more.

Specific Requirement- For the analysis we need two pivot table of bike incoming and bike outgoing count.

Function- SUM, MAX, VLOOKUP.

Formula- $T=S+E$

(S- count of bike incoming; E- count of bike outgoing; T- sum of S and E count)

Analysis Result- 3005

Visualization-

Busiest station	7th & Flower
-----------------	--------------

3. Find out the least busy station

General Description- In this analysis we have analyzed the data to find the least busy stations, least busy means where bike incoming and bike outgoing is least.

Specific Requirement- For the analysis we need two pivot table of bike incoming and bike outgoing count.

Function- SUM, MIN, VLOOKUP.

Formula- $T=S+E$

(**S**- count of bike incoming; **E**- count of bike outgoing; **T**- sum of **S** and **E** count)

Analysis Result- 3053

Visualization-

Least Busy Station 12th & Grand

4. Analyze the passholder types

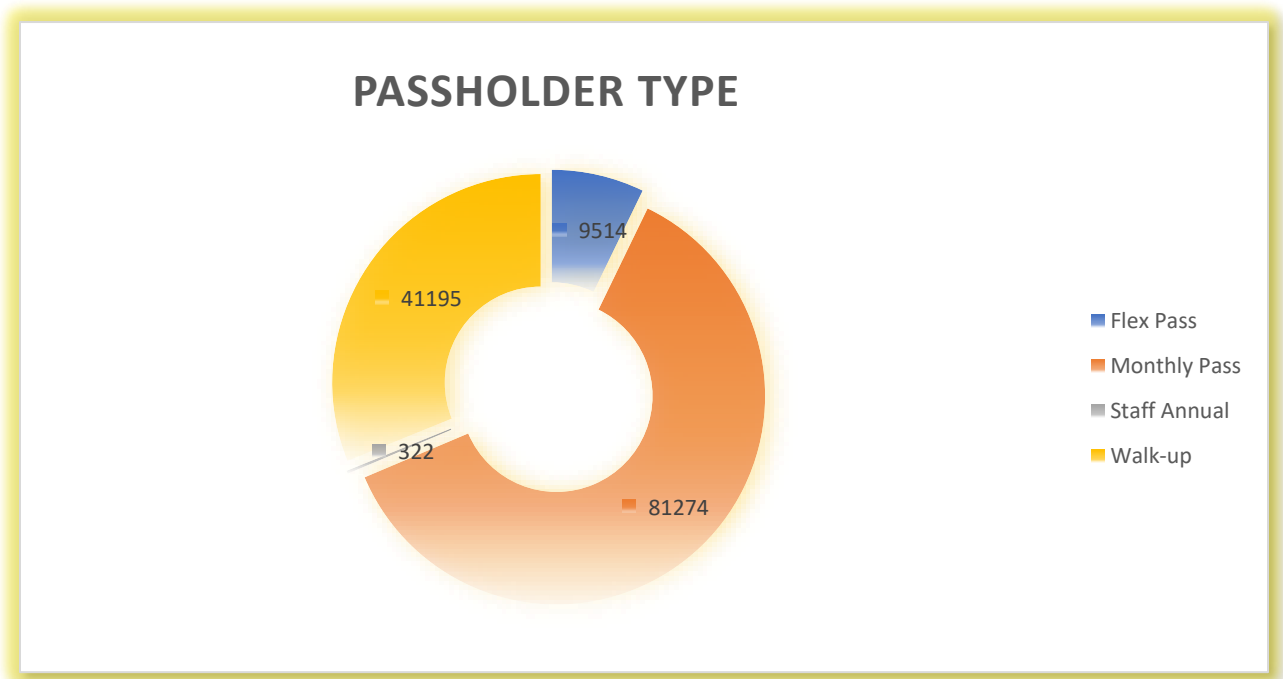
General Description- In this analysis we have analyzed the data and plot bar graph to find out the number of different types of passholder.

Specific Requirement- For the analysis we need a pivot table of passholder count.

Analysis Result-

Flex Pass	9514
Monthly Pass	81274
Staff Annual	322
Walk-up	41195

Visualization-



5. Analyze the trip category

General Description- In this analysis we have analyzed the data and plot bar graph between one-way trip and round trip.

Specific Requirement- For the analysis we need a pivot table of trip category count.

Analysis Result-

One Way	119640
Round Trip	12777

Visualization-



6. Analyze the bike condition

General Description- In this analysis we have analyzed the data and plot bar graph between bike condition is good or bad, if it is bad then need service to that bike.

Specific Requirement- For the analysis we need a pivot table between bike id and sum of duration of bikes and then make another pivot table between condition of and count of bike whose condition “Good” or “Bad”.

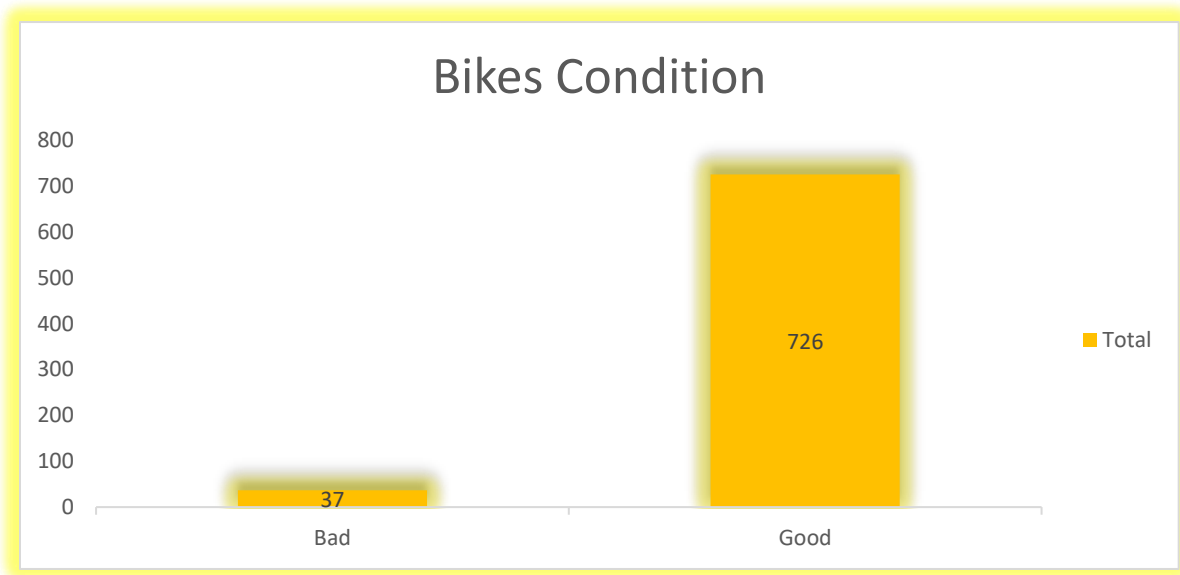
Function- IF.

Formula- IF(G4>=126,"Bad","Good").

Analysis Result-

Bad	37
Good	726

Visualization-



7. Bike information search box

General Description- In this analysis we have analyzed the data of bike using bike id, data we are analyzing like- total running time of bike, last trip id of bike, last trip date, status of bike this the bike need a service or not.

Specific Requirement- For the analysis we need data of trip id, trip date, trip duration.

Function- VLOOKUP, SUMIFS, MAXIFS, IF.

Formula- SUMIFS(duration,bike_id,"="&R55)/3600

MAXIFS(trip,bike_id,"="&R55)

VLOOKUP(R58,'metro-bike-share-trip-data'!B2:F1048576,5,FALSE)

IF(R57>=126,"Need Service","Good")

Analysis Result-

Result	
Total Travel Time:	146.37
Last Trip Id:	23794218
Last Trip Date:	2017-03-31T21:02:00
Status:	Need Service

Visualization-

Bike Information	
Enter Bike Id	4727
Result	
Total Travel Time:	146.37
Last Trip Id:	23794218
Last Trip Date:	2017-03-31T21:02:00
Status:	Need Service

8. Trip Information Search box

General Description- In this analysis we have analyzed the data of trip using trip id, data we are analyzing like- bike used in trip, trip category, passholder type, starting time and ending time of trip, source and destination of station.

Specific Requirement- - For the analysis we need data of bike id, trip date, trip duration, trip category, timing of trip, source and destination.

Function- VLOOKUP.

Formula- VLOOKUP(R64,'metro-bike-share-trip-data'!B2:C1048576,2,FALSE)

VLOOKUP(R64,'metro-bike-share-trip-data'!B2:M1048576,12,FALSE)

VLOOKUP(R64,'metro-bike-share-trip-data'!B2:N1048576,13,FALSE)

VLOOKUP(R64,'metro-bike-share-trip-data'!B2:M1048576,5,FALSE)

VLOOKUP(R64,'metro-bike-share-trip-data'!B2:M1048576,7,FALSE)

VLOOKUP(R64,'metro-bike-share-trip-data'!B:P,14,FALSE)

VLOOKUP(R64,'metro-bike-share-trip-data'!B:P,15,FALSE)

Analysis Result-

Result	
Bike Id:	6102
Trip Category:	One Way
Pass Type:	Flex Pass
Start Time:	2016-07-07T13:21:00
End Time:	2016-07-07T13:31:00
Source Station:	Wilshire & Witmer
Destination Station:	Olive & 8th

Visualization-

Trip Summary	
Enter Trip Id	1944032
Result	
Bike Id:	6102
Trip Category:	One Way
Pass Type:	Flex Pass
Start Time:	2016-07-07T13:21:00
End Time:	2016-07-07T13:31:00
Source Station:	Wilshire & Witmer
Destination Station:	Olive & 8th

9. Monthly Category wise trip analysis

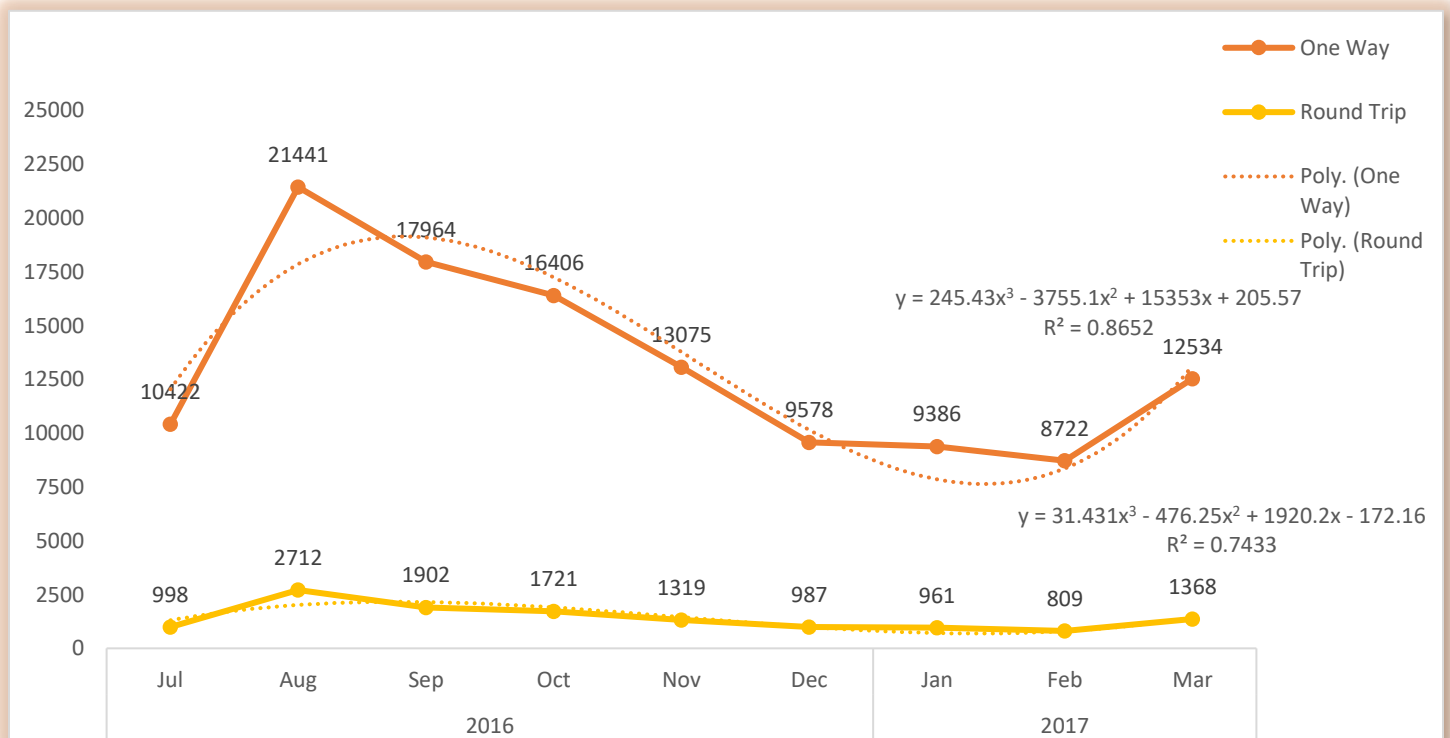
General Description- In this analysis we have analyzed the comparison between one-way or round trip on monthly basis and did prediction using trendline.

Specific Requirement- For the analysis we need the pivot table of one-way or round trip on monthly and yearly basis.

Analysis Result-

Row Labels	One Way	Round Trip
2016		
Jul	10422	998
Aug	21441	2712
Sep	17964	1902
Oct	16406	1721
Nov	13075	1319
Dec	9578	987
2017		
Jan	9386	961
Feb	8722	809
Mar	12534	1368

Visualization-

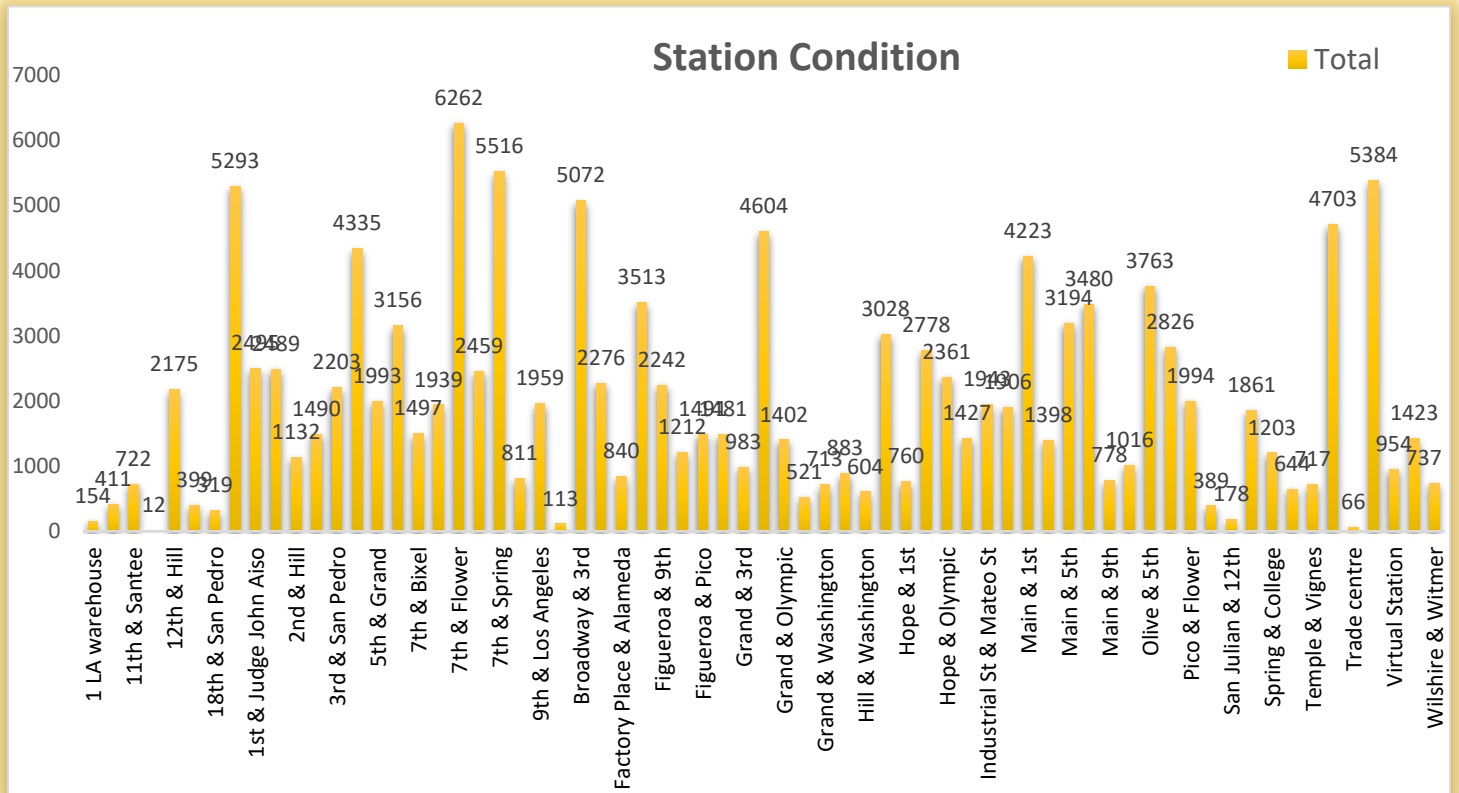


10. Station Condition

General Description- In this analysis we have analyzed the condition of every station, how much bikes are there. We are trying to solve the unbalanced station problem with this station.

Specific Requirement- For this analysis we need the pivot table between station and the count of bikes id at ending station.

Visualization-



Future Scope

The proposed system can be made more useful in future for bicycle usage. Bicycle-sharing has received increasing attention in recent years and has goal to increase cycle usage improve the connection to other modes of transit and lessen the environmental impacts of our transport activities. Also, we can provide the Gmail notification or the message notifications to the user. So, in future this work can be done.

References

<https://www.youtube.com>

<https://www.fastcompany.com/3040950/the-future-of-bike-sharing-becoming-part-of-public-transit>

<https://bikeshare.metro.net/about/data/>

Bibliography

1. Excel 2010 Bible
2. Microsoft Excel 2010 Data Analysis and Business Modeling