

Government College of Engineering Keonjhar

Department of Computer Science and Engineering

Project Title: UBER DATA ANALYSIS

BY USING PYTHON FOR DATA ANALYSIS

SL. No	Regd. No.	Name
01	1921104019	ABINASH JENA
02	1921104021	ASHUTOSH MALCHA
03	1801104139	MOHIT SAMAL
04	1921104032	SOUMYA PRAJNA BEHERA

Guided By:

Ms. Sasmita Panigrahi

PROGRESS UPDATE

SLNO.

DATE

SIGN

01

We decided to do Online RTO management system by using C#.Net and MVC. By this website we can conduct online exam for LL and vehicle registration, DL apply, renew etc:

27/09/2021

02

College had assigned Miss Sasmita Panigrahi as project guide.

04/10/2021

03

We all team mates and our guide decided to do a new project with latest technology that UBER Data Analysis - By using Python for Data Analysis

05/10/2021

04

We were new to this field that's why we had researched about Data analysis in brief, like Why Data Analysis? How it works? Why we need this? etc.

06/10/2021

Schedule & PROGRESS UPDATE

05

We gave the project abstract presentation through online mode of class.

DATE
04/10/2021

SIGN

06

Our project was approved

10/10/2021

SD

07

We checked the feasibilities of our project like market Feasibility, financial feasibility, Technical Feasibility, Operation Feasibility. From these we concluded that the Project is Feasible.

12/10/2021

SD

08

We gave the Project first Review through online mode of class.

01/11/2021

SD

SL NO	PROGRESS UPDATE	DATE	SIGN
09	we studied about Python from various source like youtube and Python documentation	01/11/2021	
10	we studied about Advance Python from youtube and That also talk about project progress with own guid	05/11/2021	
11	We studied Business problem import and export data on Python, Analyze data on python from a online course by IBM.	10/11/2021	
12	We studied about Data Wrangling, like identify and handle missing value understand data formating apply normalization to database	11/11/2021	

PROGRESS UPDATE

DATE

6/10/24

13

We collected the dataset which will we use in our project. Which we have to analyze. We collect it from a website. The data set from Network Citi.

12/11/2024

14

We tried to import Out dataset in Jupyter notebook and tried to show our set in table but we got an error.

13/11/2024

15

When we compiled the code we got result of Syntax error that (unicode error) codec can't decode bytes in position 2-3: truncated \Vxxxxxx escape. After so many research we found that.

13/11/2024

16

We did mistake while loading the dataset, we can backslash instead of forward slash. Then it worked and we moved forward.

13/11/2024

DATE	PROGRESS UPDATE	SIGN
17	<p>We studied about exploratory data analysis, Compute basic descriptive statistics create data groups. Define correlation as the linear association between two numerical variable and Define the association between two categorical variable.</p>	14/11/2021
18	<p>We tried to do the lab practical for the following with the dataset which we have studied On that day</p>	14/11/2021
19	<p>We attended the project class on offline mode and that day mate our guide and signed our project diary by her and Project diary was checked by Dr. Omkar Patankar Sir</p>	15/11/2021
20	<p>We have made some mistake in our project diary that We have listed python study from a day to another day in one table so, sir told us to differentiate that all in other table</p>	15/11/2021

date	Progress	Update	Date	Sign
21	We did another meeting with teammates and talked to guide and told about diary error and we rectify it all		16/11/2021	PL
22	We tried to attempt project work in Jupyter Notebook in first time but it shown some error in Dataset Connection.		17/11/2021	PL
23	We watched 2-3 youtube video googled about that error and tried so much thing to rectify error for connecting with Dataset in Jupyter Notebook		20/11/2021	PL
24	We skip that error for while because we think that it will take more time so we skip that and try VS code for temporarily		21/11/2021	PL

No	PROGRESS UPDATE	DATE	sign
25	we tried all that thing which we have did in previous in our course lab. we tried this time in vs code and it worked as well.	22/11/2021	PS
26	We meet to our project guide related about our project progress related to completion of our project.	23/11/2021	S
27	We did our question in vs code (first) in temporary but that time also we got an error like "permission error"	24/11/2021	S
28	We tried to rectify last error according to stack overflow We changed our dataset location and restart the vs code and once again copy and paste the new path of dataset and it worked as well.	24/11/2021	S

PROJECT PROGRESS

DATE:

sign

* Finally we choose some question to do with UBER dataset which will help full to UBER for business growth. That Question are listed below:-

1. Show the last 10 records of the dataset.
2. Show the first 10 record of the dataset.
3. Show the dimensions of the dataset.
4. Show the size of the dataset.
5. Print the information about all the variables of the dataset.
6. Check for missing value.
7. How many missing values are present.
8. Get the initial data (name it df) with dropping the 'NA' values.
9. Get summary of the original data (before dropping 'NA' value).
10. Check the information of the new data frame.

29

gr No

Project Progress

DATE

SUGR

30

11. Get the unique start destination
12. What are total Number of Unique Start destination ?
13. Print the total number of unique stop destination ?
14. Print the Uber trips that has the starting point of San Francisco
15. What is the most popular starting point for the Uber drivers ?
16. What is the most popular dropping point for the Uber drivers.
17. List of the most frequent route taken by Uber drivers 27/11/2021
18. Print all types of purposes for trip in an array
19. Plot a bar graph of purposes vs distance .
20. Print a dataframe of purpose and the distance travelled for that particular purpose .
21. plot number of trips vs category of trips
22. What is proportion of trip that is Business and What is the proportion of trips that is personal