A

Project Report on

**AMAZON ALEXA REVIEWS SENTIMENT ANALYSIS**

Submitted in partial fulfillment of completion of the course

Advanced Diploma in IT, Networking and Cloud

Submitted by:

**Roshni Verma**

Under Guidance of:

### **Ms. Arpita Roy (Edunet)**

|  |  |  |
| --- | --- | --- |
| IBM-Logo - Chicago Innovation | DGT introduces high end diploma courses - digitalLEARNING Magazine | Edunet Foundation-Delhi- CSR Organization profile |

Year 2022-2024

Abstract

Acknowledgement

Team Composition and Workload Division

Table of Contents

1. Introduction to Problem
2. Literature Review
3. Proposed Solution
4. Requirements

4.1 Technology Stack

4.2 Hardware

4.3 Software

4.4 Deployment Environment

5. User Requirements

6. Design Documentation

7. Implementation Details

8. Testing

9. Deployment

10. Future Scope

11. Conclusion

Appendix A Project Code

Appendix B Screenshot of Project

Appendix C abbreviation

References

### **Abstract**.

Amazon Alexa is a cloud-based voice service developed by Amazon that allows customers to interact with technology.   
  
There are currently over 40 million Alexa users around the world, so analyzing user sentiments about Alexa will be a good data science project.  
  
 So, if you want to learn how to analyze the sentiments of users using Amazon Alexa, this article is for you. In this article, I’ll walk you through the task of Amazon Alexa Reviews Sentiment Analysis Using Python.

**Acknowledgement**

Ms. Arpita Roy for her guidance and supervision which has provided a lot of resources needed in completing our project.

Introduction to Problem

Amazon Alexa is a cloud-based voice service developed by Amazon that allows customers to interact with technology.   
  
There are currently over 40 million Alexa users around the world, so analyzing user sentiments about Alexa will be a good data scienceproject.  
  
 So, if you want to learn how to analyze the sentiments of users using Amazon Alexa, this article is for you. In this article, I’ll walk you through the task of Amazon Alexa Reviews Sentiment Analysis Using Python.

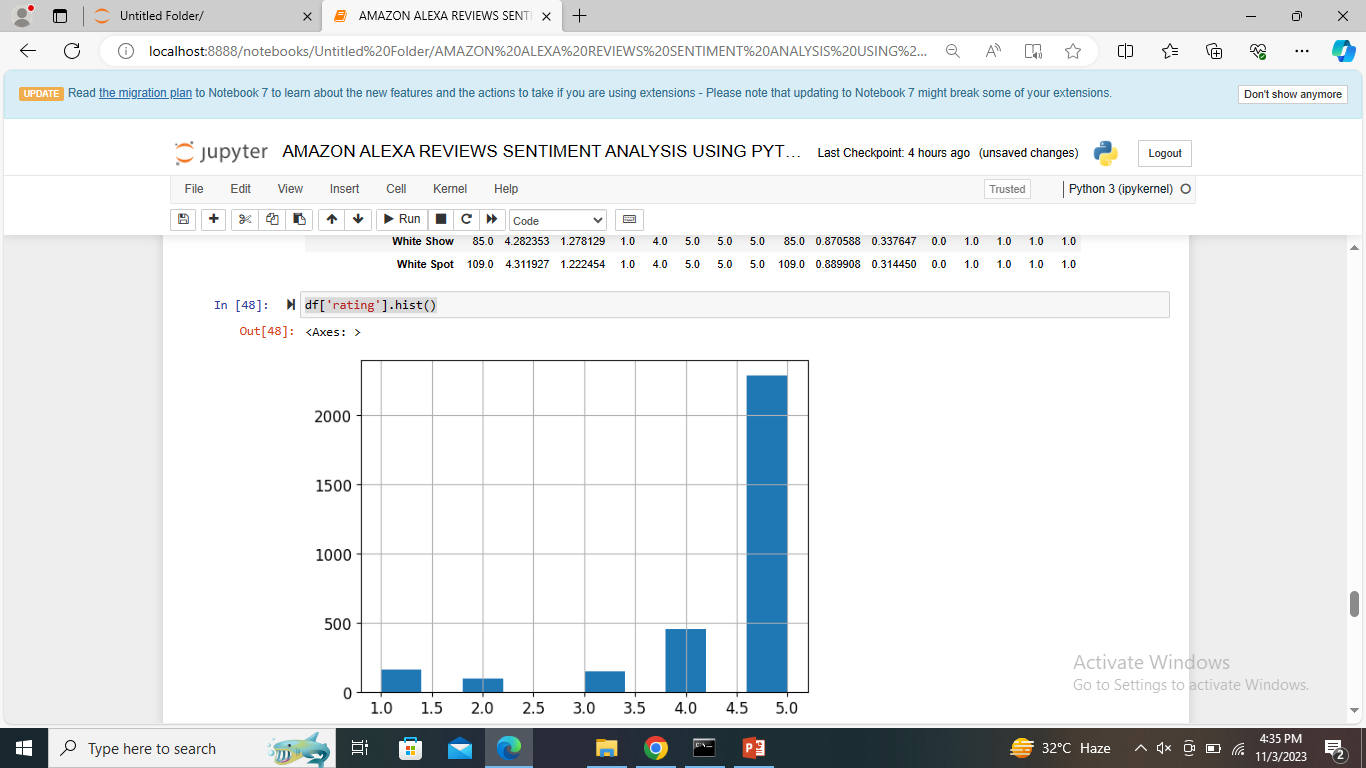
Proposed Solution

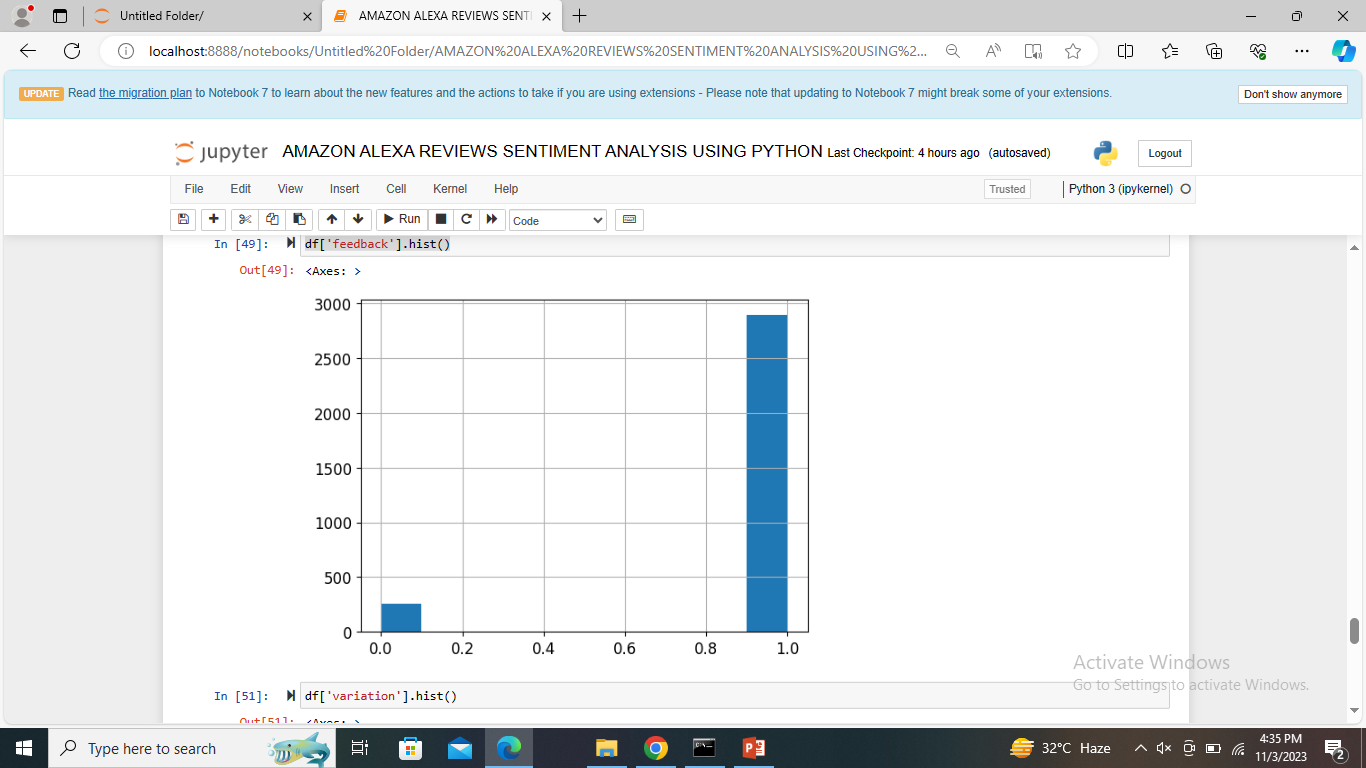
* Our goal is to check the Amazon Alexa Reviews Statement. As per my csv file I have to take the feedback ,verified reviews of people that is good or not.
* so we started with jupyter notebook and gathered data, filtered it, then moved on to visualization, and eventually created a prediction model for this data set.

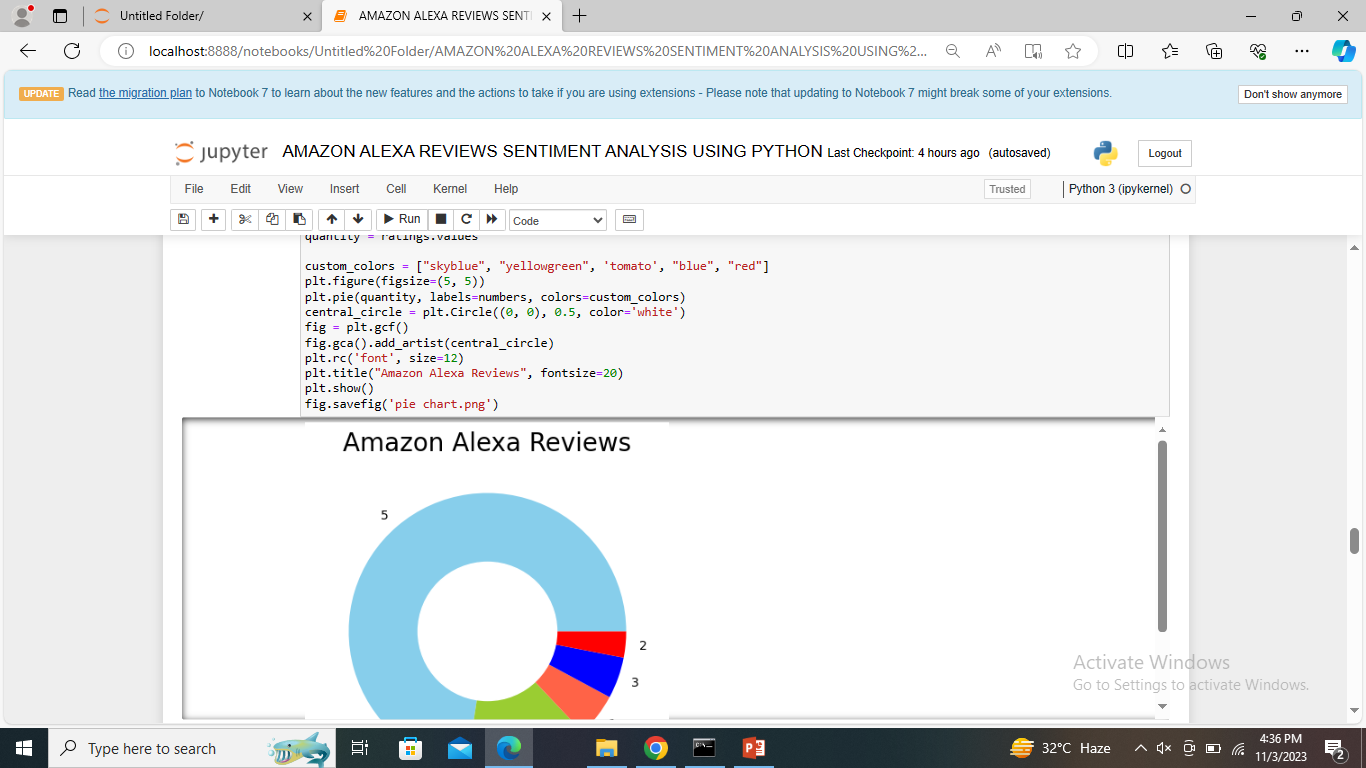
Requirements

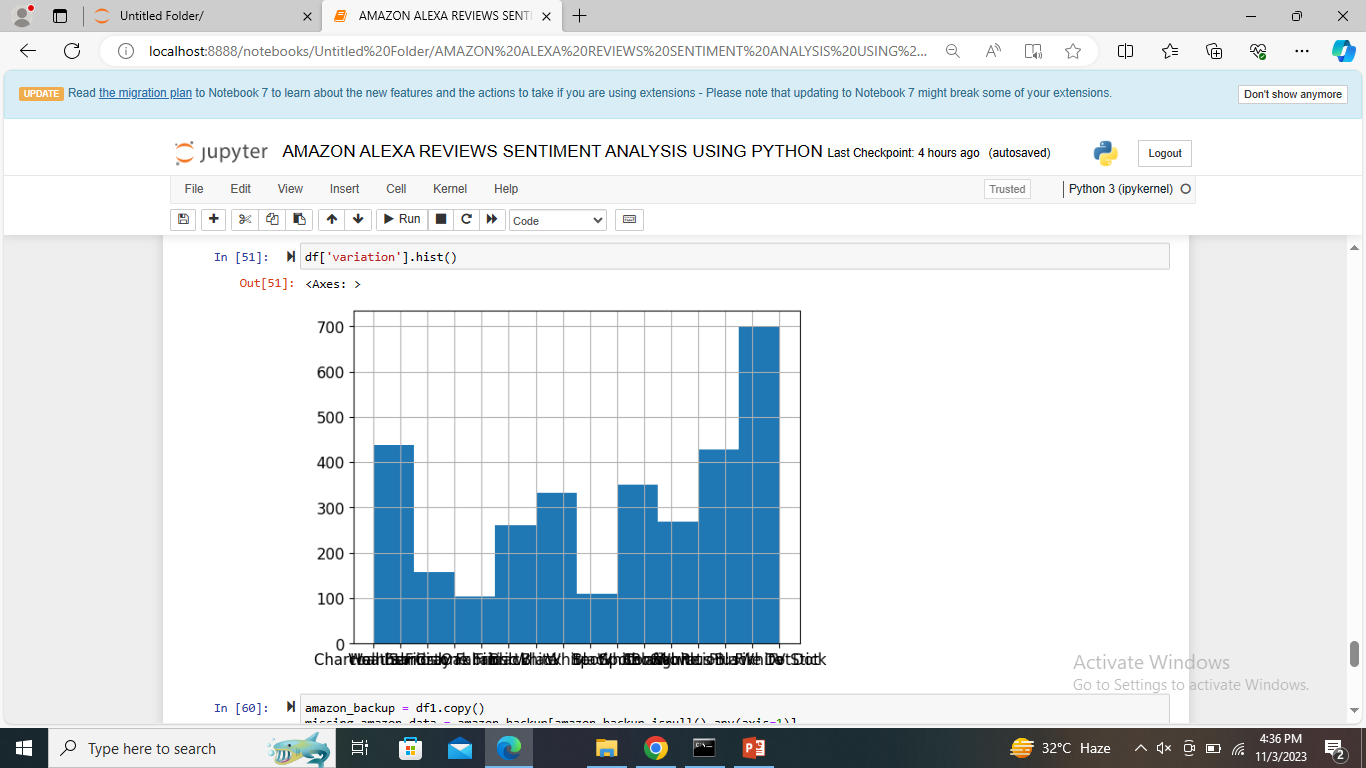
|  |  |
| --- | --- |
| Technology Stack: | PYTHON, NUMPY, PANDAS, MATHPLOTLIB, SEABORN |
| Hardware: | Processor: Intel(R) Core (TM) i5-9500 CPU @ 3.00GHz 3.00 GHz System type 64-bit operating system, x64-based processor |
| Software: | JUPYTER NOTEBOOK |
| Deployment Environment: |  |
|  |  |
|  |  |

Design Documentation









Implementation Details

**STEP1**- Firstly, I have to download csv file from Kaggle.

**STEP 2-** Secondly, I have to exact file.

**STEP 3-** I have to open JUPYTER NOTEBOOK.

**STEP 4 -** I have to import all libraries like NumPy, pandas, matplot, seaborn etc.

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

**STEP 5 -** Then, I have to read the whole csv file by this command [df = pd. read\_csv (‘AMAZON ALEXA REVIEW.csv')

print(df)]

**STEP 6 -** Then I have to give full information about csv file by using this command (df.info ()).

**STEP 7-** print(df.describe)

**STEP 8- print(df.head())#first five rows from the dataset**

**STEP 9-** **print(df.shape)#number of rows and columns**

**STEP10- print(df.columns)#displays the column names from the dataset**

**STEP11-** **print(df.dtypes)#print the datatypes of each columns**

**STEP12-** **df.info()# to get more information about your data**

**STEP13-** **print(df.tail(n=10))# get the last row**

**STEP14-** **print(df.isnull())**

**STEP 15- df.isnull().sum()#in only verified\_reviews i have only one null value**

**STEP16-** **df.isna().sum()**

**STEP17-** **#use .any() function to get concise info**

**df.isna().any()**

**STEP18-** **#. isna()detects missing values**

**df.isna()**

**STEP 19-** **df.values# values atttributes returns a numpy representation of the given DataFrame**

**STEP20-** **df1 = pd.read\_csv('output.csv', index\_col=0, na\_values=["??","????"])**

**print(df1)**

**STEP 21- ratings = df["rating"].value\_counts()**

**numbers = ratings.index**

**quantity = ratings.values**

**custom\_colors = ["skyblue", "yellowgreen", 'tomato', "blue", "red"]**

**plt.figure(figsize=(5, 5))**

**plt.pie(quantity, labels=numbers, colors=custom\_colors)**

**central\_circle = plt.Circle((0, 0), 0.5, color='white')**

**fig = plt.gcf()**

**fig.gca().add\_artist(central\_circle)**

**plt.rc('font', size=12)**

**plt.title("Amazon Alexa Reviews", fontsize=20)**

**plt.show()**

**fig.savefig('pie chart.png')**

**STEP 22-** **import numpy as np**

**print(np.var([1,9,5,6,8,7]))**

**print(np.var([4,-11,-5,16,5,7,9]))**

**STEP 23- #Describe data according to variation**

**df.groupby('variation').describe()**

**STEP 24-** **df['rating'].hist()**

**STEP 25 -** **df['feedback'].hist()**

**STEP 26- df['variation'].hist()**

**STEP 27-** **amazon\_backup = df1.copy()**

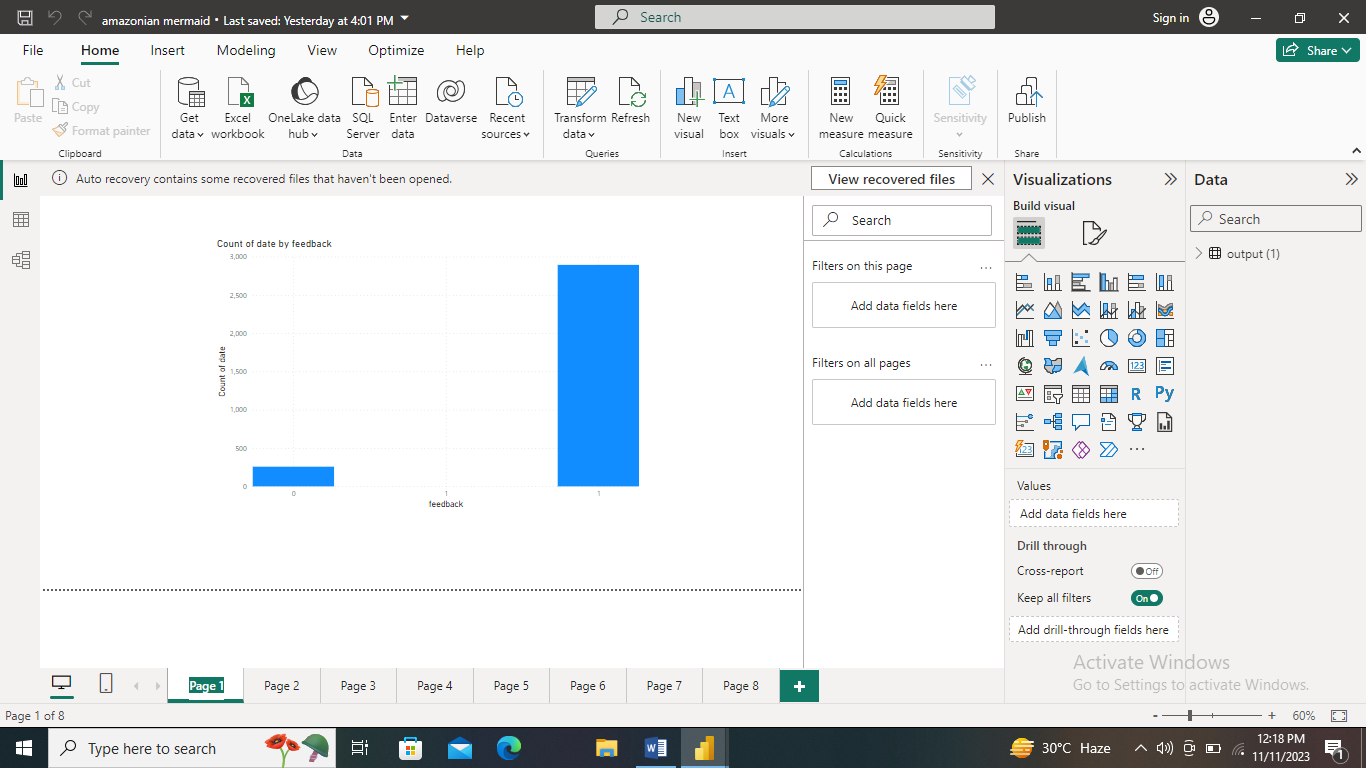
**missing\_amazon\_data = amazon\_backup[amazon\_backup.isnull().any(axis=1)]**

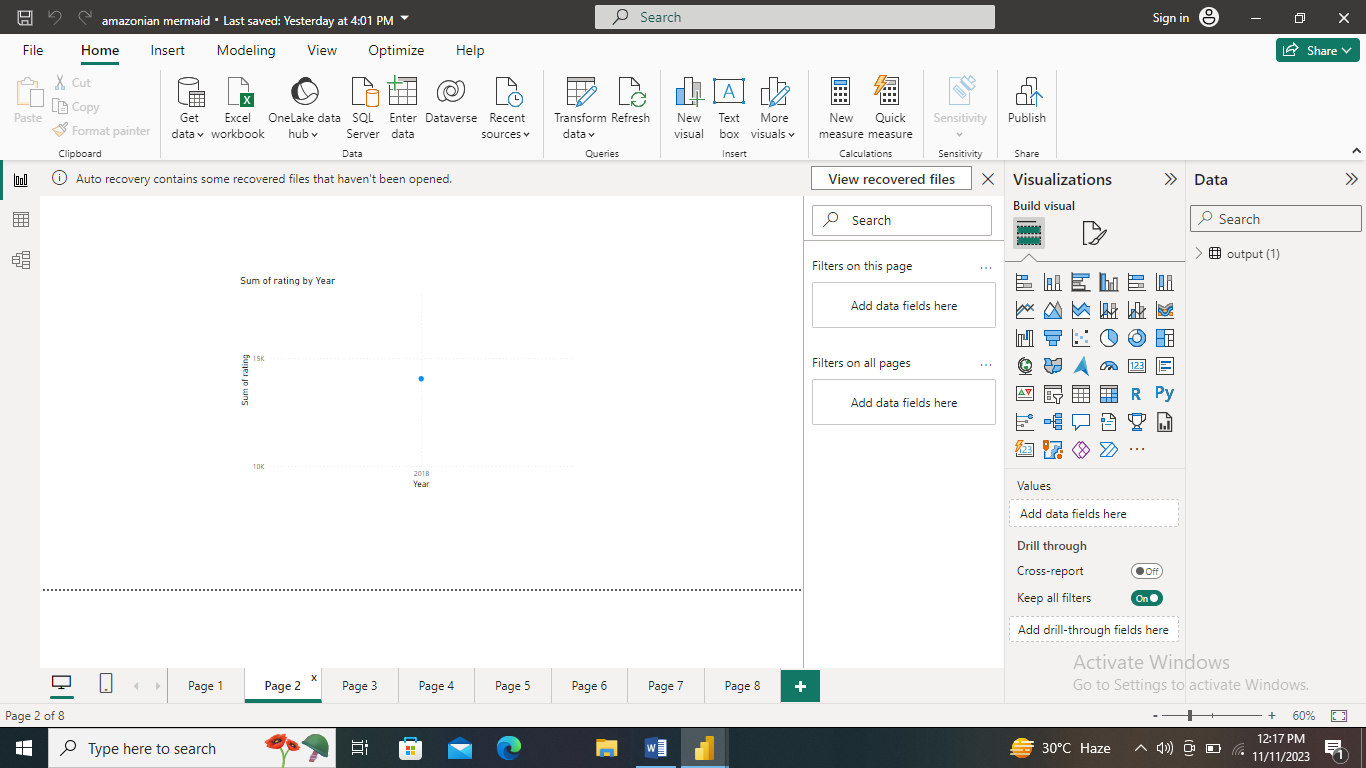
**missing\_amazon\_data**

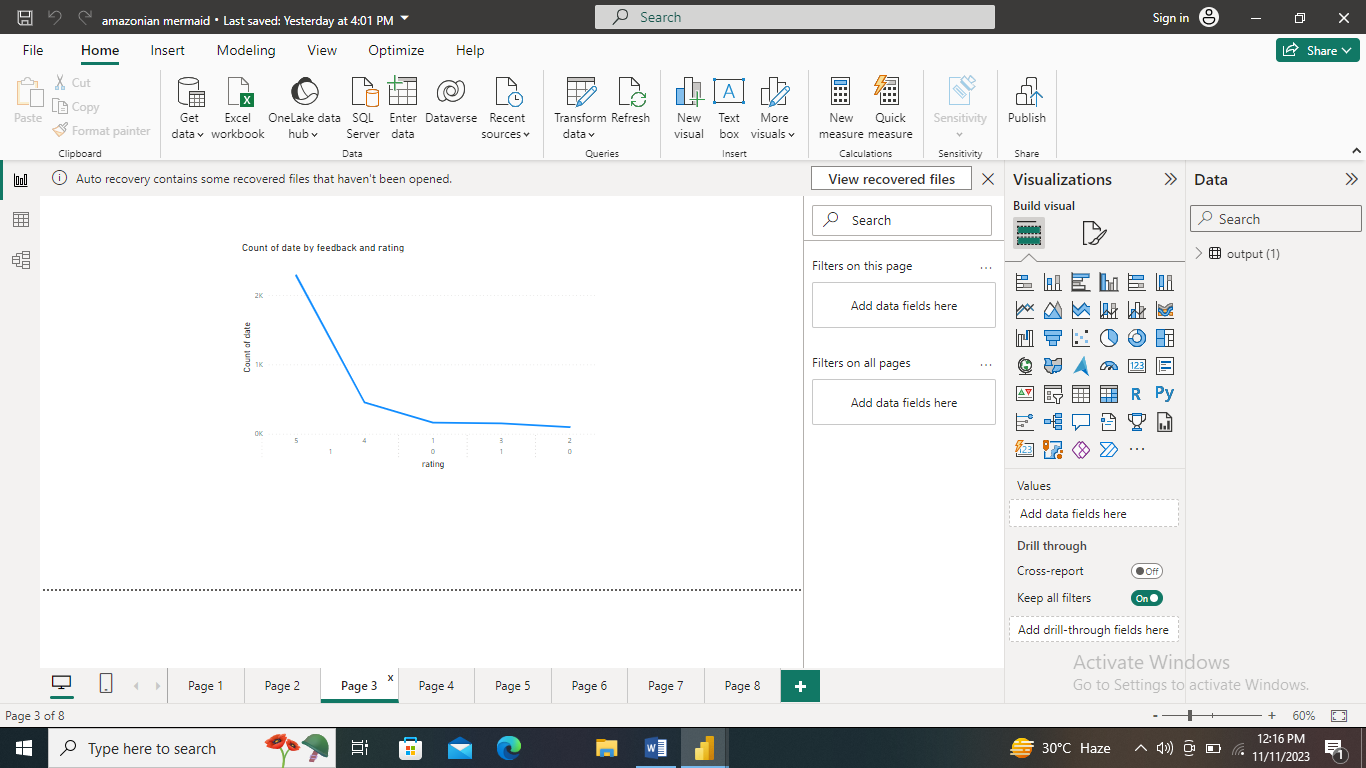
**STEP 28-** **amazon\_backup.describe()**

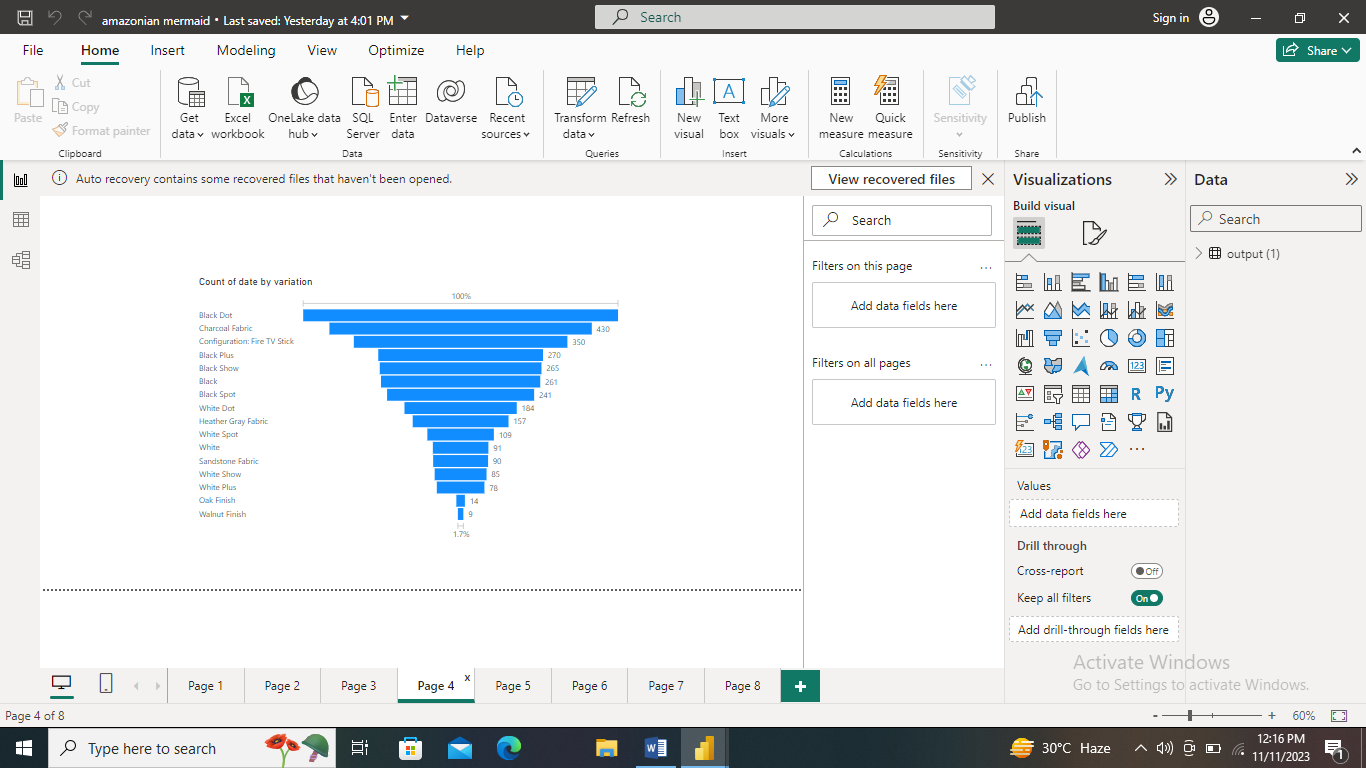
**Appendix A**

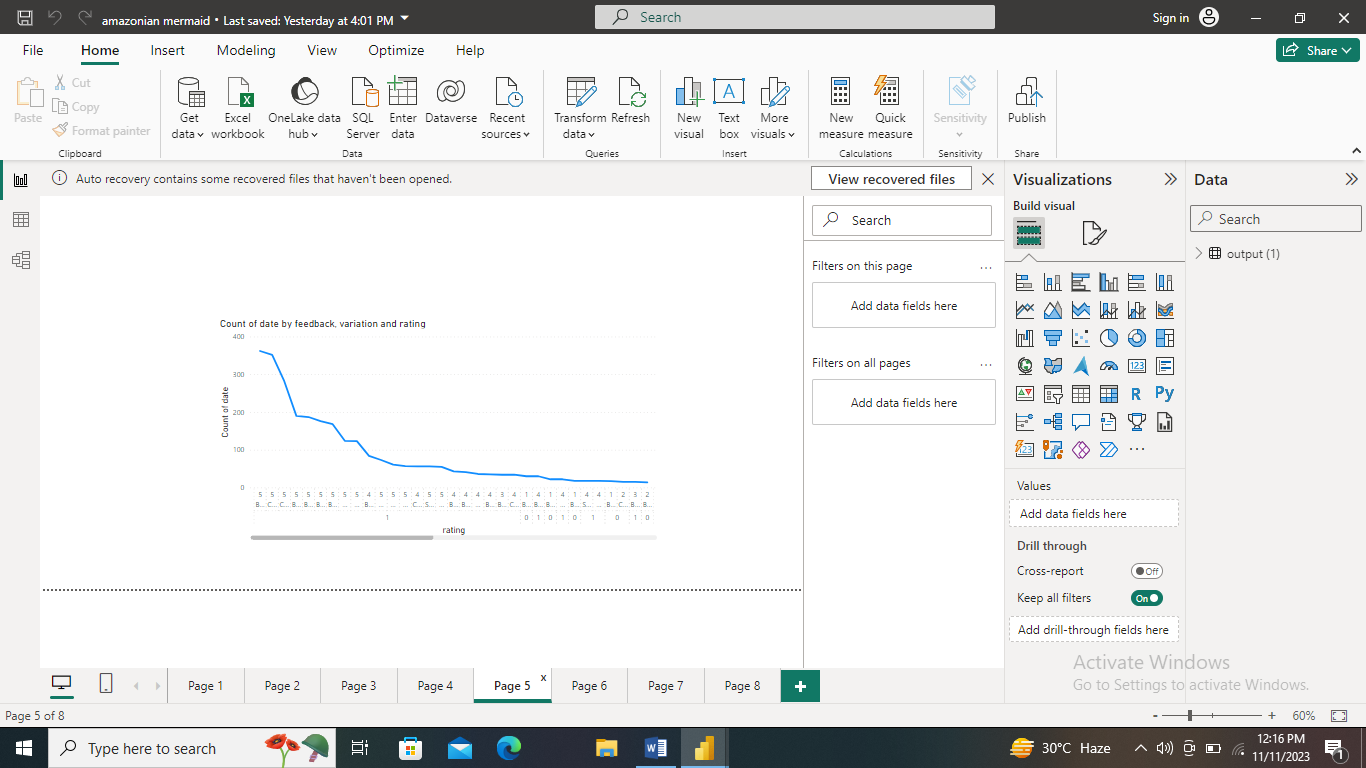
**(Project Code)**

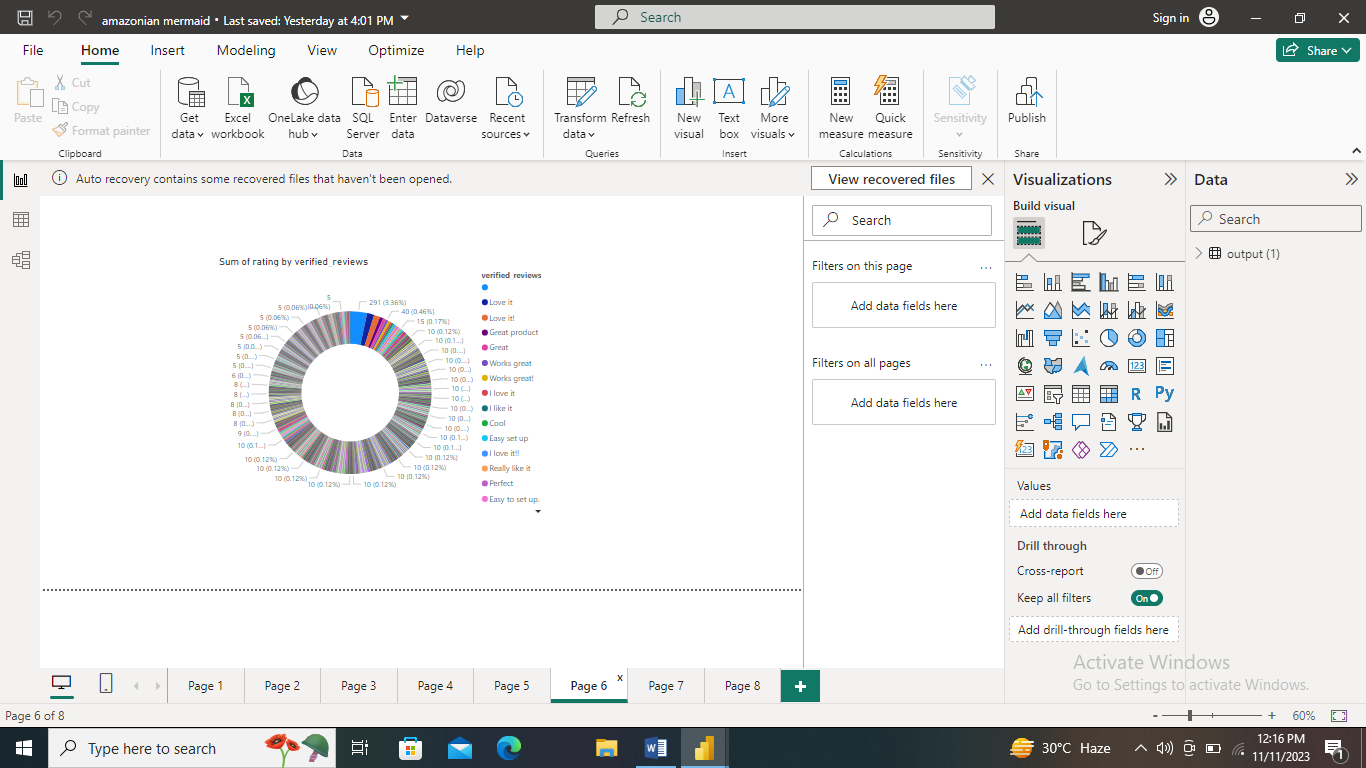
****

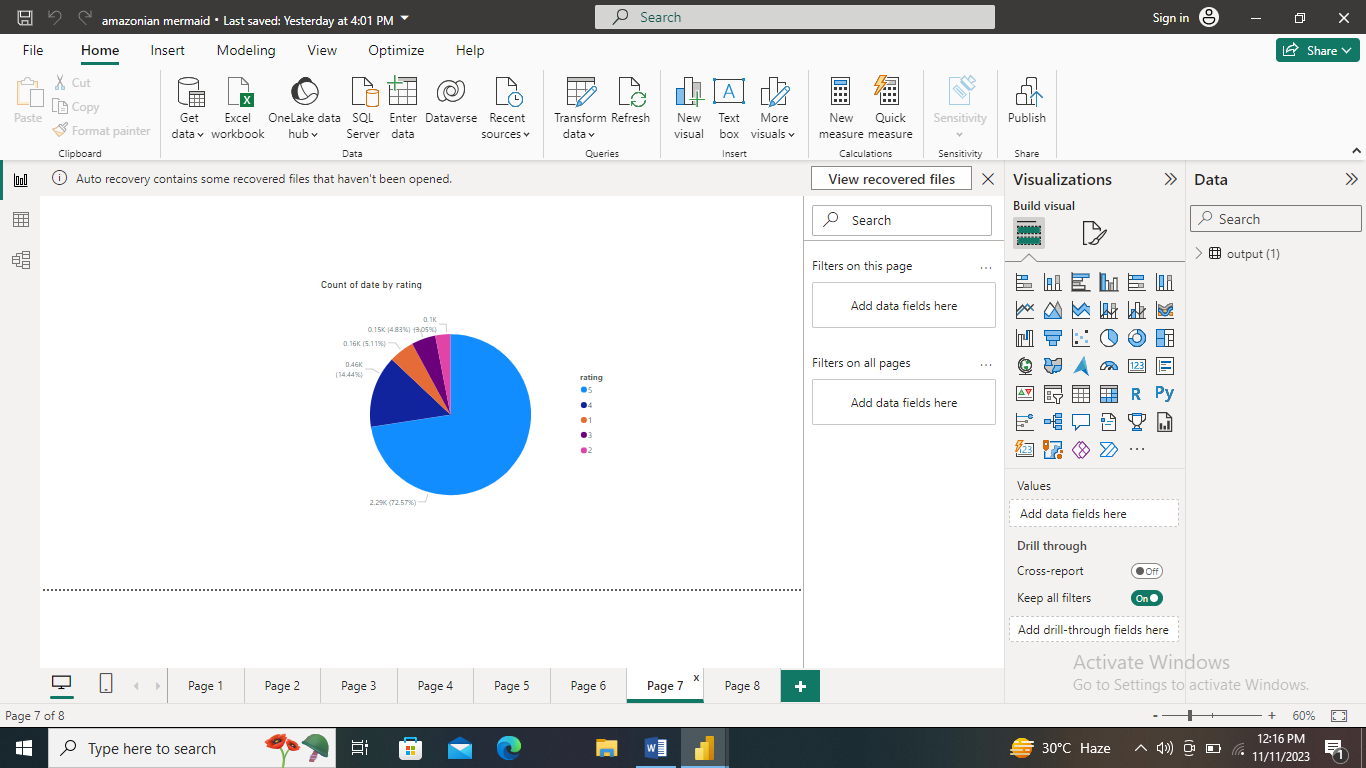
****

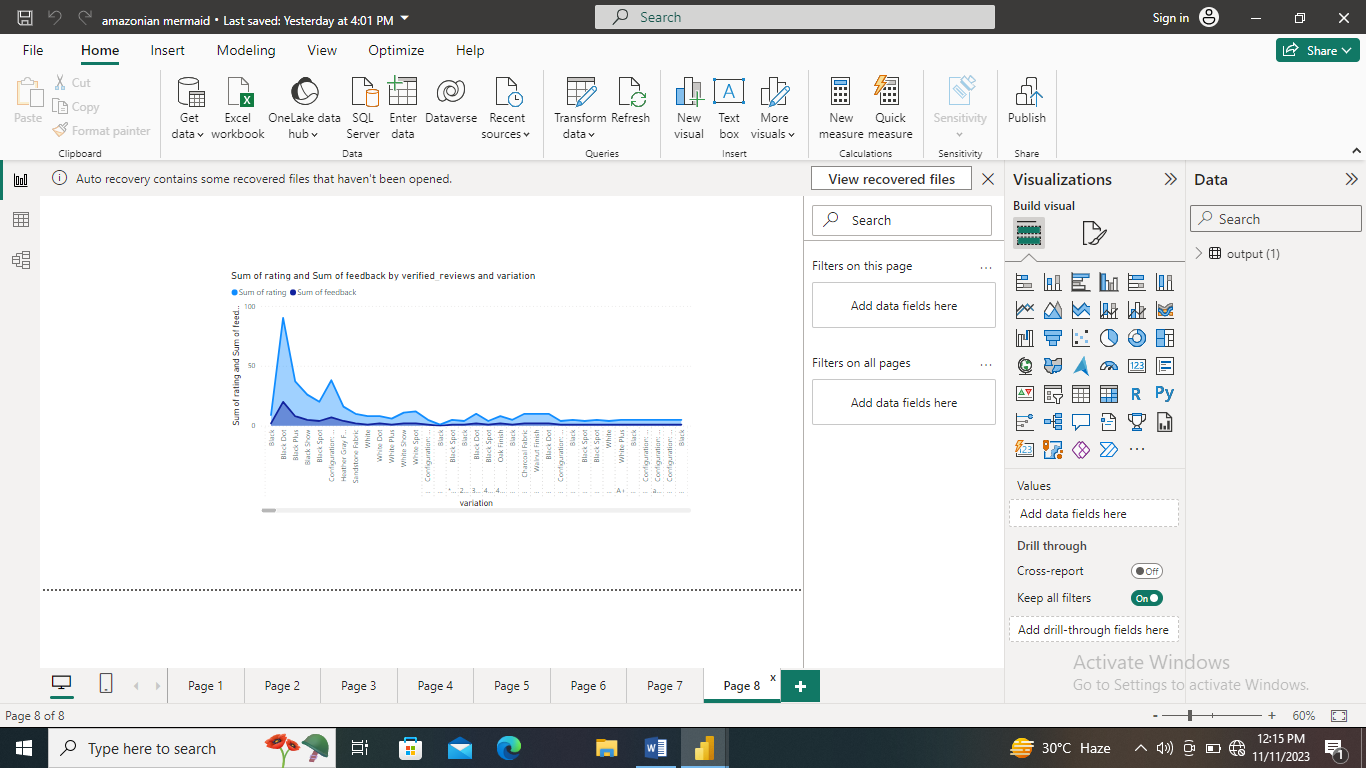
****

****

****

****

****

****

Thank You