UNIVERSITY of GREENWICH
Alliance with FPT Education

BTEC

Higher Nationals in Computing

UNIT 14

UNIT 14: BUSINESS INTELLIGENCE

ASSIGNMENT No.2

Learner's name: Trinh Thi Dieu Huyen

ID: GDD18606

Class: GCS0801_NX

Subject code: 1641

Assessor name: **NGUYEN NGOC TU**

Assignment due: March 2021

Assignment submitted: March 2021

ASSIGNMENT 2 BRIEF

Student Name/ID Number	Trinh Thi Dieu Huyen / GDD18606
Unit Number and Title	**14: Business Intelligence**
Academic Year	2020 - 2021
Unit Tutor	Nguyen Ngoc Tu
Assignment Title	**Assignment 2: Apply BI tools & techniques and their impact**
Issue Date	**March 2021**
Submission Date	March 2021
IV Name & Date	Nguyen Ngoc Tu March 2021

Submission Format
Part I: Project submission. This should be a zip / rar folder of your project, including all necessary files to run your project. There should be a link to your Tableau work on Tableau Public cloud. Part II: The submission is in the form of a group written report. This should be written in a concise, formal business style using single spacing and font size 12. You are required to make use of headings, paragraphs and subsections as appropriate, and all work must be supported with research and referenced using the Harvard referencing system. Please also provide a bibliography using the Harvard referencing system. Part III: Team needs to present their point of view about how business intelligence tools can contribute to effective decision-making as well as the legal issues involved in exploiting user data for business intelligence. You may need to research for specific examples of organizations that use BI tools to enhance or improve their business and evaluate how they can use BI tools for extend their target audience and make them more competitive within the market.

Unit Learning Outcomes
LO3 Demonstrate the use of business intelligence tools and technologies

Assignment Brief
(Continued from previous scenario) Your next task is to demonstrate to the board of directors about the ability of applying business intelligence in the company's current business processes. To demonstrate BI, you need to prepare a presentation about BI and related tools & techniques and a demonstration on real company dataset. For the presentation, you need: - Explain general concept of what is BI - Introduction to some tools / techniques for BI and their application in general For the demonstration, you need: - A (some) data set(s) extracted from the company's business processes. Explain the dataset. - Show how you pre-process data for later analysis, explain each step and it purpose - Design dashboards to show your analysis on pre-processed data. Explain clearly purpose of dashboards and charts. Suggestions should be made after analysis During the demonstration, you need collect feed-back and comments from users to review how well your dashboards design meet user or business requirement and what customization needed for future use. Team needs to present their point of view about how business intelligence tools can contribute to effective decision-making as well as the legal issues involved in exploiting user data for business intelligence. You may need to research for specific examples of organizations that use BI tools to enhance or improve their business and evaluate how they can use BI tools for extend their target audience and make them more competitive within the market. To summary, you need to submit a report in PDF includes 4 parts: your presentation, result of demonstration and review of user feedback, point of view on BI contribution and legal issues.

Learning Outcomes and Assessment Criteria		
Pass	**Merit**	**Distinction**
LO3 Demonstrate the use of business intelligence tools and technologies		**D3** Provide a critical review of the design in terms of how it meets a specific user or business requirement and identify what customisation has been integrated into the design.
P3 Determine, with examples, what business intelligence is and the tools and techniques associated with it. **P4** Design a business intelligence tool, application or interface that can perform a specific task to support problem-solving or decision-making at an advanced level.	**M3** Customise the design to ensure that it is user friendly and has a functional interface.	
LO4 Discuss the impact of business intelligence tools and technologies for effective decision-making purposes and the legal/regulatory context in which they are used		**D4** Evaluate how organisations could use business intelligence to extend their target audience and make them more competitive within the market, taking security legislation into consideration
P5 Discuss how business intelligence tools can contribute to effective decision-making. **P6** Explore the legal issues involved in the secure exploitation of business intelligence tools	**M4** Conduct research to identify specific examples of organisations that have used business intelligence tools to enhance or improve operations.	

Table of Contents

Table of Figures

Table of Tables

ASSIGNMENT 2 ANSWERS

1. Business Intelligence

1.1. Definition

Business intelligence is a wide category of applications and technologies for collecting, analyzing, and providing access to data to assist corporate users in making informed business decisions (Ranjan, 2009). For example, companies can use business intelligence to identify their most profitable customers and the underlying factors for their loyalty and identify future customers with similar, if not greater, potential. Or determine why customers leave for competitors and/or become customers using turnover and churn analysis.

1.2. Techniques

There are various techniques currently applied in business intelligence. The table below summarizes the techniques and what they mean.

Table 1: Summarizes the techniques and what they mean

TECHNIQUE	DESCRIPTION
Characterization and descriptive data mining	
Association, correlation, causality analysis (Link Analysis)	Identify relationships between attributes
Clustering and outlier analysis	Partition a set into classes, whereby items with similar characteristics are grouped together
Classification	Determine to which class a data item belongs
Exploratory Data Analysis (EDA)	Explores a data set without a strong dependence on assumptions or models; goal is to identify patterns in an exploratory manner
Model Visualization	Making discovered knowledge easily understood using charts, plots, histograms, and other visual means

1.3. Tools

There are a lot of different tools that can be used in the aid of business intelligence. These tools have been developed to help speed up the process of doing data analysis. These tools can also be summarized with a table.

Table 2: Tools

TOOL	DESCRIPTION
NUMPY	A Python library that includes a multidimensional array object, various derived objects (such as masked arrays and matrices), and a variety of routines for performing fast array and mathematical operations (Numpy, 2021).
MATPLOTLIB	A Python library that allows the creation of static, animated, and interactive visualizations (Matplotlib, 2021).
SCIKIT-LEARN	"A simple and efficient tools for predictive data analysis". Contain multiple machine learning algorithms free of implementation.
POWER BI	***Power BI is a collection of software services, applications, and connectors that work together to transform disparate data sources into logical, visually engaging, and interactive insights (Microsoft, 2021).***

1.4. Some companies that have used business Intelligence tools

Business intelligence has been designed for organizations to implement ideas for improving the operation of the company. Bellow is the name of some companies that have used business Intelligence tools to improve and enhance the operations are:

1. *Looker*
2. *Infragistics*
3. *Qlik*
4. *SAP*
5. *Microsoft*
6. *IBM*
7. *Infor*
8. *GoodData*
9. *iDashboards*
10. *Chartio*
11. *Attivio*
12. *Oracle business intelligence*
13. *Sisense*
14. *Tableau Software*
15. *Yellowfin*
16. *ThoughtSpot*
17. *Pentaho*

2. Our dataset

As previously mentioned, our dataset is about the reviews of customers for books purchased on the Amazon Kindle Store. The dataset was collected through the business processes of placing orders for the book and taking reviews. When roughly summarized, we obtain this dataset. There are a total of 9 columns for this dataset which are (not accounting for the unnamed column):

Table 3: Our dataset

asin	*The ID of the product*
helpful	*The helpfulness rating of the review – for example [2, 3] means 2/3*
overall	*The overall rating of the product*
reviewText	*The review text*
reviewTime	*The time of the review*
reviewerID	*The ID of the reviewer*
reviewerName	*The name of the reviewer*
summary	*The summary of the review*
unixReviewTime	*The unix timestamp*

As can be seen, the dataset is very raw and contains unnecessary data. So we must perform any pre-processing on the data before applying the actual analysis. We aim to perform sentiment analysis on the dataset. With sentiment analysis, organizations of all kinds may benefit from social media listening because it allows them to better understand their consumers' complaints and concerns, enabling them to scale up their services (upGrad, 2021). Due to the shift to online services, sentiment analysis assists marketers in resolving their consumers' issues or concerns. In turn, making better decisions as SA can help:

- ✓ *To enhance their services by approaching particular individuals.*
- ✓ *Keeping track of customer emotions overtime.*
- ✓ *Determine which part of the customer feels strongly about one brand.*

There are a handful of techniques for pre-processing the data, we will perform analysis after pre-processing it.

2.1. Sentiment Analysis with codes

For the coding section, we will use Jupyter Notebooks. This acts as a web-based interactive development environment to share documents, live codes, and visualization. This eases the process of running explicit codes through a separate console as the Jupyter Notebook can execute codes directly inside it. To start off, we will need to read the data into the memory, this can be done in the following codes.

```
In [8]: import numpy as np
        import pandas as pd
        import matplotlib.pyplot as plt

        df = pd.read_csv('./kindle_reviews.csv')
        df.head()
```

Out[8]:

	Unnamed: 0	asin	helpful	overall	reviewText	reviewTime	reviewerID	reviewerName	summary	unixReviewTime
0	0	B000F83SZQ	[0, 0]	5	I enjoy vintage books and movies so I enjoyed ...	05 5, 2014	A1F6404F1VG29J	Avidreader	Nice vintage story	1399248000
1	1	B000F83SZQ	[2, 2]	4	This book is a reissue of an old one; the auth...	01 6, 2014	AN0N05A9LIJEQ	critters	Different...	1388966400
2	2	B000F83SZQ	[2, 2]	4	This was a fairly interesting read. It had ol...	04 4, 2014	A795DMNCJILA6	dot	Oldie	1396569600
3	3	B000F83SZQ	[1, 1]	5	I'd never read any of the Amy Brewster mysteri...	02 19, 2014	A1FV0SX13TWVXQ	Elaine H. Turley "Montana Songbird"	I really liked it.	1392768000
4	4	B000F83SZQ	[0, 1]	4	If you like period pieces - clothing, lingo, y...	03 19, 2014	A3SPTOKDG7WBLN	Father Dowling Fan	Period Mystery	1395187200

Figure 1: Sentiment Analysis with codes

```
In [4]: df.shape
Out[4]: (982619, 10)
```

Figure 2: Sentiment Analysis with codes

The codes import necessary libraries and read the “CSV'' formatted dataset file with the panda's library, which returns a “data frame”. Then the data frame was called with the head() method which returns the first 5 rows of the dataset. And call access shape attribute to show the dimension of the data, which is nearly 1 million rows and 10 columns (9 without accounting for the unnamed one).

```
In [15]: df.isnull().sum()

Out[15]: Unnamed: 0          0
         asin                0
         helpful             0
         overall             0
         reviewText         22
         reviewTime          0
         reviewerID          0
         reviewerName     3816
         summary             1
         unixReviewTime      0
         dtype: int64
```

Figure 3: Sentiment Analysis with codes

22 rows have no reviewText, so we will remove that and remove unnecessary columns such as the ID of the product or the reviewer name/ID. Then call the head() method again to view the remaining columns.

```
In [14]: #Removing null "reviewText" rows and drop unneccessary columns
df.dropna(subset = ['reviewText'], inplace = True)
df.drop(columns = ['Unnamed: 0', 'asin', 'reviewerID', 'reviewerName', 'summary', 'unixReviewTime'], inplace = True)
df.reviewTime = pd.to_datetime(df.reviewTime)
df.head()
```

Out[14]:

	helpful	overall	reviewText	reviewTime
0	[0, 0]	5	I enjoy vintage books and movies so I enjoyed ...	2014-05-05
1	[2, 2]	4	This book is a reissue of an old one; the auth...	2014-01-06
2	[2, 2]	4	This was a fairly interesting read. It had ol...	2014-04-04
3	[1, 1]	5	I'd never read any of the Amy Brewster mysteri...	2014-02-19
4	[0, 1]	4	If you like period pieces - clothing, lingo, y...	2014-03-19

Figure 4: Sentiment Analysis with codes

We can try to plot the number of reviews per year. This might indicate how the website has become more popular, or the people trend in reading more books.

```
In [29]: ax = df.reviewTime.dt.year.value_counts().sort_index().plot(kind = 'bar')

         plt.title('Number of reviews per year')
         plt.xlabel('Year')
         plt.ylabel('Number of reviews')
         plt.show()
```

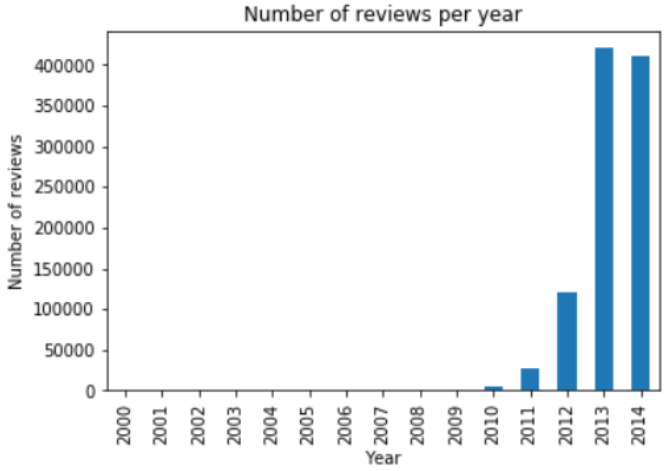

Figure 5: Sentiment Analysis with codes

In order to see the relation between these remaining columns, we can look for the correlation between these features. But we can only do that by converting all columns to numerical values. This can be done as follows.

```
In [106]: df['helpful'] = df['helpful'].apply(lambda x: eval(x))

In [143]: def helpfulToScore(r, perBookReviews):
              asin, helpful = r
              return ( helpful[0] / helpful[1] ) / perBookReviews[asin] if (helpful[0] <= helpful[1] and helpful[1] != 0) else helpful[1]

In [146]: perBookReviews = df['asin'].value_counts().to_dict()
          print(list(perBookReviews.items())[0:5])
          df['helpful'] = df[['asin', 'helpful']].apply(lambda r: helpfulToScore(r, perBookReviews) , axis = 1)
          df.head()

          [('B006GWO5WK', 1112), ('B00BTIDW4S', 781), ('B00BT0J8ZS', 516), ('B00JDYC5OI', 502), ('B00H0V069M', 481)]
```

Out[146]:

	asin	helpful	overall	reviewText	reviewTime
0	B000F83SZQ	0.000	5	I enjoy vintage books and movies so I enjoyed ...	2014-05-05
1	B000F83SZQ	0.125	4	This book is a reissue of an old one; the auth...	2014-01-06
2	B000F83SZQ	0.125	4	This was a fairly interesting read. It had ol...	2014-04-04
3	B000F83SZQ	0.125	5	I'd never read any of the Amy Brewster mysteri...	2014-02-19
4	B000F83SZQ	0.000	4	If you like period pieces - clothing, lingo, y...	2014-03-19

Figure 6: Sentiment Analysis with codes

Since the “helpful” column is a string, we need to convert it to an array by calling eval(). Then to convert that array to a “numerical value”, we might need to define some method. We can do this by taking the average helpfulness, we also need to normalize it by dividing it by the total number of reviews for that specific book.

```
In [169]: sw = set(stopwords.words('english'))
          stemmer = PorterStemmer()
          lemmatizer = WordNetLemmatizer()

          df['reviewText'] = df['reviewText'] = df['reviewText'].apply(lambda x: x.lower())

In [170]: df['reviewText'] = df['reviewText'].apply(lambda x: x.translate(str.maketrans('', '', string.punctuation)))

In [171]: df['reviewText'] = df['reviewText'].apply(lambda x: ' '.join(x.split()))

In [172]: df['reviewText'] = df['reviewText'].apply(lambda x: ' '.join([word for word in word_tokenize(x) if word not in sw]))

In [173]: df['reviewText'] = df['reviewText'].apply(lambda x: ' '.join([lemmatizer.lemmatize(word, pos = 'v') for word in word_tokenize(x)]

In [174]: df.head()
```

Out[174]:

	asin	helpful	overall	reviewText	reviewTime
0	B000F83SZQ	0.000	5	enjoy vintage book movies enjoy read book plot...	2014-05-05
1	B000F83SZQ	0.125	4	book reissue old one author bear 1910 era say ...	2014-01-06
2	B000F83SZQ	0.125	4	fairly interest read old style terminologyi gl...	2014-04-04
3	B000F83SZQ	0.125	5	id never read amy brewster mysteries one reall...	2014-02-19
4	B000F83SZQ	0.000	4	like period piece clothe lingo enjoy mystery a...	2014-03-19

Figure 7: Sentiment Analysis with codes

We need to normalize or pre-process the text first by removing punctuation characters, white space, removing stopwords, and lemmatizing the text. Since we need to plot the correlation, we should copy the original dataframe to a new one so it would not be affected. And then convert reviewText to the total number of words and reviewTime to its timestamp.

```
In [178]: df2 = df.copy()

In [210]: df2['reviewText'] = df2['reviewText'].apply(lambda x: len(word_tokenize(x)))

In [213]: df2['reviewTime'] = df2['reviewTime'].values.astype(np.int64)

In [215]: df2.drop(columns = 'asin', inplace = True)

In [216]: df2.head()
```

Out[216]:

	helpful	overall	reviewText	reviewTime
0	0.000	5	29	1399248000000000000
1	0.125	4	45	1388966400000000000
2	0.125	4	31	1396569600000000000
3	0.125	5	9	1392768000000000000
4	0.000	4	12	1395187200000000000

Figure 8: Sentiment Analysis with codes

Now we can show the correlation between these columns.

```
In [225]: plt.subplots(figsize=(8,5))
          sn.heatmap(df2.corr())
```

```
Out[225]: <matplotlib.axes._subplots.AxesSubplot at 0x1e4fe51f288>
```

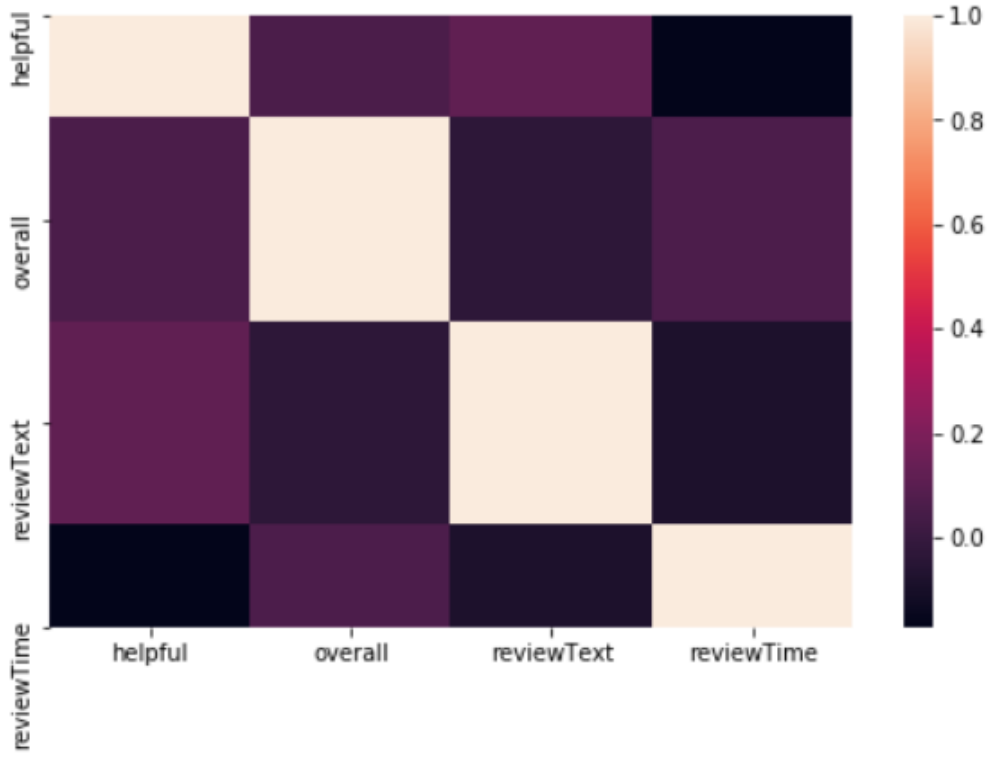

Figure 9: Sentiment Analysis with codes

As we can see, values that are near 0 indicate little-to-no-relationship between the features. So we can assume that these columns are independent of one another. This time we will remove the reviewTime and the helpful columns. We retain only the reviewText column as it is the most related to sentiment analysis.

```
In [162]: df3 = df['overall']

In [163]: df3.shape
Out[163]: (982597,)

In [164]: df3.head()
Out[164]: 0    5
          1    4
          2    4
          3    5
          4    4
          Name: overall, dtype: int64

In [208]: df4 = df3.apply(lambda x: 2 if x == 5 else 1 if 2 <= x <= 4 else 0)

In [202]: df4.value_counts()
Out[202]: 2    829256
          1    130323
          0     23018
          Name: overall, dtype: int64
```

Figure 10: Sentiment Analysis with codes

Now we will create labels by making overall values equal 1 as “bad”, between 2 and 4 inclusive as “okay”, and equal 5 as “good”. As the data needs to be numbered instead of the string so we will make “bad” as 0, “okay” as 1, “good” as 2.

```
In [209]: df5 = df['reviewText'][df4.index]
          df5.head()
Out[209]: 0    enjoy vintage book movies enjoy read book plot...
          1    book reissue old one author bear era say nero ...
          2    fairly interest read old style terminologyi gl...
          3    id never read amy brewster mysteries one reall...
          4    like period piece clothe lingo enjoy mystery a...
          Name: reviewText, dtype: object
```

Figure 11: Sentiment Analysis with codes

Now, df5 contains reviews and df4 contains the labels. We will use Multinomial Naïve Bayes for classifying these reviews into their corresponding “sentiment”.

```
In [204]: x_train, x_test, y_train, y_test = train_test_split(df5 , df4 , random_state = 0)

          print(x_train.head())
          print(y_train.head())
```

```
512473    story delightful great story children trouble ...
61111     read book one book two series week book use na...
30000     book sound interest purcased blind normally wa...
875667    much say raw read ruby great character felt fe...
443254    nice quick summer read story character great i...
Name: reviewText, dtype: object
512473    2
61111     1
30000     2
875667    1
443254    1
Name: overall, dtype: int64
```

Figure 12: Sentiment Analysis with codes

```
In [199]: c = TfidfVectorizer()
          c_train = c.fit_transform(x_train)
          c_test = c.transform(x_test)
```

```
In [210]: nb = MultinomialNB()
          nb.fit(c_train, y_train_small)
          print('Training Accuracy: {:.2f}'.format(nb.score(c_train, y_train)))
          print('Testing Accuracy: {:.2f}'.format(nb.score(c_test, y_test)))
```

```
Training Accuracy: 0.84
Testing Accuracy: 0.84
```

Figure 13: Sentiment Analysis with codes

We then perform training on the “x_train” dataset and then use it to predict the new dataset. The accuracy score is shown below. As we can see, predicting new data gives 84% accuracy.

2.2. Plotting in Power BI

For the application section, we will use Power BI. The application can be used much easier with just plotting data. As pre-processing with Power BI can be extremely difficult if one is not familiar with DAX syntax. We first import the data into Power BI.

Column1	asin	helpful	overall	reviewText	reviewTime	reviewerID	reviewerName	summary	unixRev
26	B000FBFMVG	[0, 0]	5	Really shouldn't have Han Solo on the cover as he doesn't	Saturday, February 15, 2014	A3AZH0OO0GUAI3	Amazon Customer	An interesting short story.	
194	B000JMLBHU	[0, 0]	5	KEEPS YOU INTERESTED. RECOMMEND TO ANY AGE. GEN	Tuesday, January 29, 2013	A1BKERTS8SPHER	Amazon Customer	THE MYSTERIOUS ISLAND	
367	B000JQUT8S	[0, 0]	5	Once again, I have finally read this classic story for the firs	Saturday, March 23, 2013	A2L8LCZ7C9XMTG	Amazon Customer	Another Great Classic	
386	B000JQUTGK	[0, 0]	5	I love reading things involving our 16th president.I got to l	Thursday, May 2, 2013	A1L6S1VV8XWFLA	Amazon Customer	Wonderful stories	
392	B000JQUU00	[0, 0]	5	I love reading fairy tales from different area of the world.It	Thursday, May 2, 2013	A1L6S1VV8XWFLA	Amazon Customer	True Fairy Tales	
462	B000NJL7Y6	[0, 0]	5	I have all of these books in the first set of Westmoreland s	Wednesday, February 26, 2014	A3W4PH6ZA5M3AI	Amazon Customer	Love	
505	B000OI11GK	[0, 0]	5	Great story. This author is really good. First time reading h	Thursday, January 3, 2013	A2C7Y9VZ9OYRT1	Amazon Customer	McKettricks Bundle	
580	B000R93D4Y	[0, 0]	5	a lot of Bianca D'Arc since I found the first paperback in a l	Sunday, August 4, 2013	A22ZJ11DIOJYXW	Amazon Customer	Have read ...	
657	B000R93DPS	[0, 0]	5	Whenever I just want a good read, I pick Stripped Bare. I'n	Thursday, November 3, 2011	A3RSU55US4Y9CW	Amazon Customer	In love with this book	
662	B000R93DQC	[0, 0]	5	The day (Friday) before AJ is set to marry Kyle, she finds ou	Tuesday, January 21, 2014	A185Y88G2ICSLD	Amazon Customer	Another great read!!!!!!!!!!	
675	B000R93E3E	[0, 0]	5	Loved the storyline between the 2 main characters. The w	Wednesday, July 17, 2013	A1R951S6YTALTO	Amazon Customer	great romance. Good read	
736	B000SWV6HG	[0, 0]	5	Lori foster has captured my attention with her stories. she	Tuesday, October 1, 2013	A34N7F4UPQV86A	Amazon Customer	great read!	
971	B000XZKZKM	[0, 0]	5	Rafe and Emma forever, I luved this book so much and wis	Monday, August 1, 2011	AZHZZONN1SO6P	Amazon Customer	LUV IT, LUV IT!!!	
1149	B0013N884S	[0, 0]	5	Each and every one of these stories keeps you on the edge	Friday, March 28, 2014	A3CWYI90CBZD9G	Amazon Customer	Another must read for the vampire genre enthusiast	
1178	B00149M19E	[0, 0]	5	I love On the Edge! Meeting Brett, Leo and Lisa through th	Wednesday, November 27, 2013	A2MZC4MYRF8WQV	Amazon Customer	Second time around and still amazing	
1302	B0015YEQNM	[0, 0]	5	I love fantasy and these books give you that. The characte	Sunday, February 24, 2013	A2U8UHEP1F0RBG	Amazon Customer	Loved these stories	
1314	B0015Z7VFG	[0, 0]	5	A beautifully written story by Maya Banks of love and forgi	Thursday, March 27, 2014	A39I9BNI1FJCME	Amazon Customer	Fantastic!	
1431	B0016IXMDW	[0, 0]	5	This is a great book that you just can't put down.... a very g	Thursday, February 21, 2013	A2CR88S46RV9SW	Amazon Customer	just friends?	
1509	B001892DFC	[0, 0]	5	Love, love, love this book. The author did a wonderful job	Tuesday, March 18, 2014	A39I9BNI1FJCME	Amazon Customer	Love It	
1540	B001892DGG	[0, 0]	5	I always know that when I buy one of her books I am gonn	Sunday, March 2, 2014	AW9H7MT8X4A6D	Amazon Customer	Another good book	
1595	B001892DJS	[0, 0]	5	I enjoyed every minute reading this book. I love Time Trav	Wednesday, March 19, 2014	A11Y8MVERD1MHU	Amazon Customer	Sexy Time Travel Romance ... YES!!!	
1611	B001892DWA	[0, 0]	5	Another great story by Maya Banks. While reading this sto	Sunday, June 1, 2014	A39I9BNI1FJCME	Amazon Customer	Enjoyed it immensely	
1659	B001892EI8	[0, 0]	5	I love mystery baby books where the daddy did not know	Thursday, January 24, 2013	A3TWXRH2DP213Y	Amazon Customer	I love this book	
1661	B001892EI8	[0, 0]	5	This story is great for reading and it shows that love betwe	Saturday, May 24, 2014	A3CVJFR3SHPJOG	Amazon Customer	Great love story between two people	

Figure 14: Plotting in Power BI

Remove unnecessary columns from the table, which can include null rows and other rows such as reviewerName, reviewerID,…

	asin	helpful	overall	reviewText	reviewTime
1	B000F83SZQ	[0, 0]	5	I enjoy vintage books and movies so I enjoyed reading this book. The ...	5/5/2014
2	B000F83SZQ	[2, 2]	4	This book is a reissue of an old one; the author was born in 1910. It's o...	1/6/2014
3	B000F83SZQ	[2, 2]	4	This was a fairly interesting read. It had old- style terminology.I was gl...	4/4/2014
4	B000F83SZQ	[1, 1]	5	I'd never read any of the Amy Brewster mysteries until this one.. So I ...	2/19/2014
5	B000F83SZQ	[0, 1]	4	If you like period pieces - clothing, lingo, you will enjoy this mystery. A...	3/19/2014
6	B000F83SZQ	[0, 0]	4	A beautiful in-depth character description makes it like a fast pacing m...	5/26/2014
7	B000F83SZQ	[0, 0]	4	I enjoyed this one tho I'm not sure why it's called An Amy Brewster My...	6/10/2014
8	B000F83SZQ	[1, 1]	4	Never heard of Amy Brewster. But I don't need to like Amy Brewster t...	3/22/2014
9	B000FA64PA	[0, 0]	5	Darth Maul working under cloak of darkness committing sabotage no...	10/11/2013
10	B000FA64PA	[0, 0]	4	This is a short story focused on Darth Maul's role in helping the Trade ...	2/13/2011
11	B000FA64PA	[0, 0]	5	I think I have this one in both book and audio. It is a good story either ...	1/27/2014
12	B000FA64PA	[0, 0]	4	Title has nothing to do with the story. I did enjoy it though. Good sho...	9/17/2011
13	B000FA64PA	[0, 0]	3	Well written. Interesting to see Sideous (through Maul) maneuvering e...	12/31/2013
14	B000FA64PK	[0, 0]	3	Troy Denning's novella Recovery was originally published in e-book for...	3/15/2012
15	B000FA64PK	[0, 0]	5	I am not for sure on how much of a difference the short story has on t...	5/12/2013
16	B000FA64PK	[0, 0]	5	I really enjoyed the book. Had the normal back against the wall mome...	1/2/2014
17	B000FA64PK	[0, 0]	5	Great read enjoyed every minute of it . I think it answers some questi...	10/29/2013
18	B000FA64PK	[4, 4]	3	Another well written eBook by Troy Denning, but why is it put where it...	4/16/2009
19	B000FA64PK	[0, 1]	5	This one promises to be another good book. I have been reading throu...	1/27/2014
20	B000FA64PK	[0, 0]	4	I have a version of "Star by Star" that does not include the novella "Re...	7/20/2012
21	B000FA64PK	[0, 0]	5	Excellent! Very well written story, very exciting with LOTS of action. Th...	12/31/2013
22	B000FA64QO	[0, 0]	2	With Ylesia, a novella originally published in e-book form, Walter Jon ...	3/15/2012
23	B000FA64QO	[0, 0]	5	Great book couldn't put it down. The story expands events in the new...	2/4/2014
24	B000FA64QO	[2, 2]	4	Most of the New Jedi Order books focus on the Solo kids, to the extent...	5/17/2011
25	B000FA64QO	[0, 0]	5	I was hoping to find this one in book form. The story looks like it will b...	1/27/2014
26	B000FA64QO	[0, 0]	3	The events of "Ylesia" take place during "Destiny's Way." If you've rea...	8/7/2012

Query Settings

PROPERTIES

Name: kindle_reviews

All Properties

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Removed Other Columns
- Removed Null Rows Of revie...

Figure 15: Plotting in Power BI

Then we can plot the number of reviews per year using this DAX expression by creating a measure.

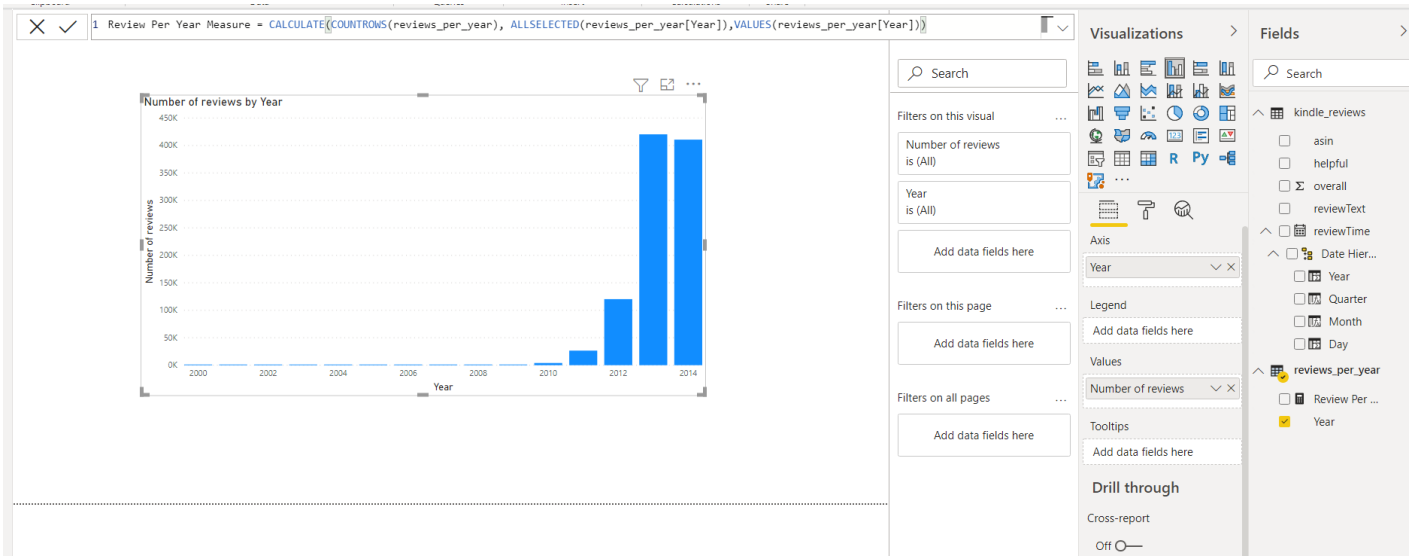

Figure 16: Plotting in Power BI

Trying to do machine learning inside an application like this is really hard as we lack the freedom we have as in code. Thought possible, but it can be quite overwhelming and time-consuming.

3. BI in effective decision-making

There are various steps, tools and techniques of business intelligence, which are used by various organizations to implement better ideas in order to achieve the goals.

- **For example**, *suppose for the hospital, companies can collect the data from the server of all the people who are trying to search for a hospital in order to spend vacations or any other reasons. After collecting the data, they can extract the information of the people who haven't booked any hospital yet and try to connect to them in order to attract them to visit the hospital. Once they visit, members of the hospital can try and build a good patient relationship which will leave a good impact on the patients and the staff will be bound to visit and admit the hospital. These are the steps that can be taken to run the business of the hospital, but to implement these things, certain decisions are supposed to be made by the junior and senior authorities of the hospital at their level.*

This is a very important aspect of the business. The decisions are also supposed to be taken by considering the security concerns as well which will make the data of the guest more secure and confidential.

Business intelligence tools help companies achieve success and customer satisfaction by presenting and displaying early decisions that ensure customer loyalty and admiration (Atwah, 2013). The decision-making process is closely linked to the business intelligence framework and these tools. The decision-making process is aided by a business intelligence system, which helps to save time, money, and human error. In the world of business, a business intelligence system is technological globalization that enables an enterprise to deal with modern intelligence revolution customs. If the BI system achieves its goals, decision-making processes will be effective, direct, and time-saving, risks will be reduced, and there will be no room for ambiguity in the workplace. Building a fraud detection model, for example, may assist decision-making processes for preventing fraud or planning audit strategies but it's important to recognize that useful outcomes aren't necessarily guaranteed in data mining models.

Business intelligence systems aid in optimizing business processes and capital, the maximization of sales, and the improvement of proactive decision-making (UE, 2006). The systems in question could be used to build a number of applications in finance, competition monitoring, accounting, marketing, and manufacturing, among other fields. In order to make successful decisions, it is important to gather information from various sources and then incorporate it.

4. Legal issues in BI

One could have all the data required to make better hiring decisions, cut costs, and run a more effective department. However, data is useless if it cannot be placed into context. Business intelligence entails not only collecting and storing data, but also viewing, analyzing, and using it to enhance operations. Business intelligence tools help to build visual reports that show the bigger picture. Therefore, to obtain data without any issues, one must also have legal concerns for BI.

- **Ethical**: *Data mining, especially data about people, has significant ethical implications. When determining whether or not to notify a person that his or her information is being processed for potential data mining, businesses face an ethical dilemma. A company's competitive advantage in the market can be harmed by allowing people to opt-out of data collection. An organization must determine if a lack of ethical consideration would result in a loss of consumer goodwill and a backlash from customers (UKEssays, 2021).*

- **Data sharing**: *Important data can only be shared with those who have a "need to know" in a business. Maintaining data protection should be a priority. Employees who need access to such data should explain why they need it and what they intend to do with it. This can possibly violate the laws in some countries if any data breach happens*.

REFERENCES

[1] J. Ranjan. 2009. Business Intelligence: Concepts, Components, Techniques and Benefits. [online] Available at: <http://www.jatit.org/volumes/research-papers/Vol9No1/9Vol9No1.pdf> [Accessed 11 March 2021].

[2] upGrad blog. 2021. *Sentiment Analysis: What is it and Why Does it Matter? | upGrad blog*. [online] Available at: <https://www.upgrad.com/blog/sentiment-analysis-what-is-it-and-why-does-it-matter/> [Accessed 12 March 2021].

[3] Numpy.org. 2021. *What is NumPy? — NumPy v1.20 Manual*. [online] Available at: <https://numpy.org/doc/stable/user/whatisnumpy.html> [Accessed 12 March 2021].

[4] Matplotlib.org. 2021. *Matplotlib: Python plotting — Matplotlib 3.3.4 documentation*. [online] Available at: <https://matplotlib.org/> [Accessed 12 March 2021].

[5] Docs.microsoft.com. 2021. *What is Power BI? - Power BI*. [online] Available at: <https://docs.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview> [Accessed 12 March 2021].

[6] M. Ahwat. 2013. [online] Available at: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.402.8414&rep=rep1&type=pdf> [Accessed 14 March 2021].

[7] University Of Economics. 2006. [online] Available at: <http://ijikm.org/Volume1/IJIKMv1p047-058Olszak19.pdf> [Accessed 14 March 2021].

[8] UKEssays.com. 2021. *Data Mining The Privacy And Legal Issues Information Technology Essay*. [online] Available at: <https://www.ukessays.com/essays/information-technology/data-

mining-the-privacy-and-legal-issues-information-technology-essay.php> [Accessed 15 March 2021].

[9] Matt David, 2020. *7 Real-World Examples of Business Intelligence in Use*. [online] Chartio. Available at: <https://chartio.com/learn/business-intelligence/7-real-world-examples-of-business-intelligence-in-use/> [Accessed 16 March 2021].

[10] Hillsberg, A., 2017. *Top 5 business intelligence (BI) solutions for your company*. [online] Big Data Made Simple. Available at: <https://bigdata-madesimple.com/top-5-business-intelligence-bi-solutions-for-your-company/> [Accessed 16 March 2021].

[11] Haije, E., 2021. *Top 15 Business Intelligence Tools: An Overview - Mopinion*. [online] Mopinion. Available at: <https://mopinion.com/business-intelligence-bi-tools-overview/> [Accessed 16 March 2021].

[12] Tableau. 2021. *5 real examples of business intelligence in action*. [online] Available at: <https://www.tableau.com/learn/articles/business-intelligence-examples> [Accessed 16 March 2021].

[13] Kaggle.com. 2021. *Amazon Top 50 Bestselling Books 2009 - 2019*. [online] Available at: <https://www.kaggle.com/sootersaalu/amazon-top-50-bestselling-books-2009-2019?fbclid=IwAR11EsAl2uhS-Xs-cdqR3WPd1-4hGxK0f4Iu4PtUsB6mjfj-ca9jAKdFYD0> [Accessed 16 March 2021].

[14] DataFlair. 2021. *Understand Power BI Filters to take your business to next level! - DataFlair*. [online] Available at: <https://data-flair.training/blogs/power-bi-filter/?fbclid=IwAR2Gm13_kE4m4zpP3ypRFmet5Zm4X-gvYqxgE7yjpaWl4rLjkXtLlxKVk6A> [Accessed 16 March 2021].